DEX-0113 B6

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
23 August 2001 (23.08.2001)

PCT

(10) International Publication Number
WO 01/61055 A2

(51) International Patent Classification[7]: C12Q 1/68

(21) International Application Number: PCT/US01/05674

(22) International Filing Date: 20 February 2001 (20.02.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/183,188 17 February 2000 (17.02.2000) US

(71) Applicant: DIADEXUS, INC. [US/US]; 3303 Octavius Drive, Santa Clara, CA 95054 (US).

(72) Inventors: CHEN, Sei-Yu; 160 Mira Street, Foster City, CA 94404 (US). SUN, Yongming; Apartment 260, 869 S. Winchester Boulevard, Apartment 260, San Jose, CA 95128 (US). MACINA, Roberto, A.; 4118 Crescendo Avenue, San Jose, CA 95136 (US).

(74) Agents: LICATA, Jane, Massey et al.; Licata & Tyrrell P.C., 66 E. Main Street, Marlton, NJ 08053 (US).

(81) Designated States *(national)*: AU, CA, JP.

(84) Designated States *(regional)*: European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 01/61055 A2

(54) Title: METHODS FOR DIAGNOSING, MONITORING, STAGING, IMAGING AND TREATING LUNG CANCER VIA LUNG CANCER SPECIFIC GENES

(57) Abstract: The invention relates to LSG polypeptides, polynucleotides encoding the polypeptides, methods for producing the polypeptides, in particular by expressing the polynucleotides, and agonists and antagonists of the polypeptides. The invention further relates to methods for utilizing such polynucleotides, polypeptides, agonists and antagonists for applications, which relate, in part, to research, diagnostic and clinical arts.

METHODS FOR DIAGNOSING, MONITORING, STAGING, IMAGING AND TREATING LUNG CANCER VIA LUNG CANCER SPECIFIC GENES

This application claims priority to U.S. Provisional Application. No. 60/183,188, filed February 17, 2000, the contents of which are incorporated herein by reference in its entirety.

FIELD OF THE INVENTION

This invention relates, in part, to newly identified polynucleotides and polypeptides; variants and derivatives of the polynucleotides and polypeptides; processes for making the polynucleotides and the polypeptides, and their variants and derivatives; agonists and antagonists of the polypeptides; and uses of the polynucleotides, polypeptides, variants, derivatives, agonists and antagonists. In particular, in these and in other regards, the invention relates to polynucleotides and polypeptides hereinafter referred to as "Lng103" and "Lng104".

BACKGROUND OF THE INVENTION

Lung cancer is the second most prevalent type of cancer for both men and women in the United States and is the most common cause of cancer death in both sexes. Lung cancer can result from a primary tumor originating in the lung or a secondary tumor which has spread from another organ such as the bowel or breast. Primary lung cancer is divided into three main types; small cell lung cancer; non-small cell lung cancer; and mesothelioma. Small cell lung cancer is also called "Oat Cell" lung cancer because the cancer cells are a distinctive oat shape. There are three types of non-small cell lung cancer. These are grouped

together because they behave in a similar way and respond to treatment differently to small cell lung cancer. The three types are squamous cell carcinoma, adenocarcinoma, and large cell carcinoma. Squamous cell cancer develops from the cells that line the airways. Adenocarcinoma also develops from the cells that line the airways. However, adenocarcinoma develops from a particular type of cell that produces mucus (phlegm). Large cell lung cancer has been thus named because the cells look large and rounded when they are viewed under a microscope. Mesothelioma is a rare type of cancer which affects the covering of the lung called the pleura. Mesothelioma is often caused by exposure to asbestos.

Secondary lung cancer is cancer that has started somewhere else in the body (for example, the breast or bowel) and spread to the lungs. Choice of treatment for secondary lung cancer depends on where the cancer started. In other words, cancer that has spread from the breast should respond to breast cancer treatments and cancer that has spread from the bowel should respond to bowel cancer treatments.

The stage of a cancer indicates how far a cancer has spread. Staging is important because treatment is often decided according to the stage of a cancer. The staging is different for non-small cell and for small cell cancers of the lung.

Non-small cell cancer can be divided into four stages. Stage I is very localized cancer with no cancer in the lymph nodes. Stage II cancer has spread to the lymph nodes at the top of the affected lung. Stage III cancer has spread near to where the cancer started. This can be to the chest wall, the covering of the lung (pleura), the middle of the chest (mediastinum) or other lymph nodes. Stage IV cancer has spread to another part of the body.

Since small cell lung cancer can spreads quite early

in development of the disease, small cell lung cancers are divided into only two groups. These are: limited disease, that is cancer that can only be seen in one lung and in nearby lymph nodes; and extensive disease, that is cancer that has spread outside the lung to the chest or to other parts of the body. Further, even if spreading is not apparent on the scans, it is likely that some cancer cells will have broken away and traveled through the bloodstream or lymph system. To be safe, it is therefore preferred to treat small cell lung cancers as if they have spread, whether or not secondary cancer is visible. Because surgery is not typically used to treat small cell cancer, except in very early cases, the staging is not as critical as it is with some other types of cancer. Chemotherapy with or without radiotherapy is often employed. The scans and tests done at first will be used later to see how well a patient is responding to treatment.

Accordingly, there is a great need for more sensitive and accurate methods for the staging of a cancer in a human to determine whether or not such cancer has metastasized and for monitoring the progress of a cancer in a human which has not metastasized for the onset of metastasis.

Two diagnostic markers have now been identified for lung cancer, and in particular squamous cell carcinoma of the lung. These diagnostic markers are referred to herein generally as lung specific genes or LSGs and more specifically as Lng103 and Lng104.

In the present invention, methods are provided for detecting, diagnosing, monitoring, staging, prognosticating, imaging and treating lung cancer via the lung specific genes referred to herein as LSGs. For purposes of the present invention, LSG refers, among other things, to native protein expressed by the gene comprising a polynucleotide sequence of SEQ ID NO:1, 2, 5 or 6. An amino acid sequence encoded by the polynucleotide of SEQ ID

NO:1 is depicted in SEQ ID NO:3. An amino acid sequence encoded by the polynucleotide of SEQ ID NO:2 is depicted in SEQ ID NO:4. By "LSG" it is also meant herein polynucleotides which, due to degeneracy in genetic coding, comprise variations in nucleotide sequence as compared to SEQ ID NO: 1, 2, 5 or 6, or, but which still encode the same protein. In the alternative, what is meant by LSG as used herein, means the native mRNA encoded by the gene comprising the polynucleotide sequence of SEQ ID NO: 1, 2, 5 or 6 levels of the gene comprising the polynucleotide sequence of SEQ ID NO: 1, 2, 5 or 6, or levels of a polynucleotide which is capable of hybridizing under stringent conditions to the antisense sequence of SEQ ID NO: 1, 2, 5 or 6.

Other objects, features, advantages and aspects of the present invention will become apparent to those of skill in the art from the following description. It should be understood, however, that the following description and the specific examples, while indicating preferred embodiments of the invention are given by way of illustration only. Various changes and modifications within the spirit and scope of the disclosed invention will become readily apparent to those skilled in the art from reading the following description and from reading the other parts of the present disclosure.

SUMMARY OF THE INVENTION

Toward these ends, and others, it is an object of the present invention to provide LSGs comprising a polynucleotide of SEQ ID NO:1, 2, 5 or 6, a protein expressed by a polynucleotide of SEQ ID NO:1, 2, 5 or 6 or variant thereof which expresses the protein; or a polynucleotide which is capable of hybridizing under stringent conditions to the antisense sequence of SEQ ID NO: 1, 2, 5 or 6. Preferred LSG proteins are depicted in

SEQ ID NO:3 and 4.

It is another object of the present invention to provide a method for diagnosing the presence of lung cancer by analyzing for changes in levels of LSG in cells, tissues or bodily fluids compared with levels of LSG in preferably the same cells, tissues, or bodily fluid type of a normal human control, wherein a change in levels of LSG in the patient versus the normal human control is associated with lung cancer.

Further provided is a method of diagnosing metastatic lung cancer in a patient having lung cancer which is not known to have metastasized by identifying a human patient suspected of having lung cancer that has metastasized; analyzing a sample of cells, tissues, or bodily fluid from such patient for LSG; comparing the LSG levels in such cells, tissues, or bodily fluid with levels of LSG in preferably the same cells, tissues, or bodily fluid type of a normal human control, wherein an increase in LSG levels in the patient versus the normal human control is associated with lung cancer which has metastasized.

Also provided by the invention is a method of staging lung cancer in a human which has such cancer by identifying a human patient having such cancer; analyzing a sample of cells, tissues, or bodily fluid from such patient for LSG; comparing LSG levels in such cells, tissues, or bodily fluid with levels of LSG in preferably the same cells, tissues, or bodily fluid type of a normal human control sample, wherein an increase in LSG levels in the patient versus the normal human control is associated with a cancer which is progressing and a decrease in the levels of LSG is associated with a cancer which is regressing or in remission.

Further provided is a method of monitoring lung cancer in a human having such cancer for the onset of metastasis. The method comprises identifying a human

patient having such cancer that is not known to have metastasized; periodically analyzing a sample of cells, tissues, or bodily fluid from such patient for LSG; comparing the LSG levels in such cells, tissue, or bodily fluid with levels of LSG in preferably the same cells, tissues, or bodily fluid type of a normal human control sample, wherein an increase in LSG levels in the patient versus the normal human control is associated with a cancer which has metastasized.

Further provided is a method of monitoring the change in stage of lung cancer in a human having such cancer by looking at levels of LSG in a human having such cancer. The method comprises identifying a human patient having such cancer; periodically analyzing a sample of cells, tissues, or bodily fluid from such patient for LSG; comparing the LSG levels in such cells, tissue, or bodily fluid with levels of LSG in preferably the same cells, tissues, or bodily fluid type of a normal human control sample, wherein an increase in LSG levels in the patient versus the normal human control is associated with a cancer which is progressing and a decrease in the levels of LSG is associated with a cancer which is regressing or in remission.

Further provided are methods of designing new therapeutic agents targeted to an LSG for use in imaging and treating lung cancer. For example, in one embodiment, therapeutic agents such as antibodies targeted against LSG or fragments of such antibodies can be used to treat, detect or image localization of LSG in a patient for the purpose of detecting or diagnosing a disease or condition. In this embodiment, an increase in the amount of labeled antibody detected as compared to normal tissue would be indicative of tumor metastases or growth. Such antibodies can be polyclonal, monoclonal, or omniclonal or prepared by molecular biology techniques. The term "antibody", as used

herein and throughout the instant specification is also meant to include aptamers and single-stranded oligonucleotides such as those derived from an *in vitro* evolution protocol referred to as SELEX and well known to those skilled in the art. Antibodies can be labeled with a variety of detectable and therapeutic labels including, but not limited to, radioisotopes and paramagnetic metals. Therapeutic agents such as small molecules and antibodies which decrease the concentration and/or activity of LSG can also be used in the treatment of diseases characterized by overexpression of LSG. Such agents can be readily identified in accordance with teachings herein.

Other objects, features, advantages and aspects of the present invention will become apparent to those of skill in the art from the following description. It should be understood, however, that the following description and the specific examples, while indicating preferred embodiments of the invention, are given by way of illustration only. Various changes and modifications within the spirit and scope of the disclosed invention will become readily apparent to those skilled in the art from reading the following description and from reading the other parts of the present disclosure.

GLOSSARY

The following illustrative explanations are provided to facilitate understanding of certain terms used frequently herein, particularly in the examples. The explanations are provided as a convenience and are not limitative of the invention.

DIGESTION of DNA refers to catalytic cleavage of the DNA with a restriction enzyme that acts only at certain sequences in the DNA. The various restriction enzymes referred to herein are commercially available and their reaction conditions, cofactors and other requirements for

use are known and routine to the skilled artisan.

For analytical purposes, typically, 1 μg of plasmid or DNA fragment is digested with about 2 units of enzyme in about 20 μl of reaction buffer. For the purpose of isolating DNA fragments for plasmid construction, typically 5 to 50 μg of DNA are digested with 20 to 250 units of enzyme in proportionately larger volumes.

Appropriate buffers and substrate amounts for particular restriction enzymes are described in standard laboratory manuals, such as those referenced below, and they are specified by commercial suppliers.

Incubation times of about 1 hour at 37°C are ordinarily used, but conditions may vary in accordance with standard procedures, the supplier's instructions and the particulars of the reaction. After digestion, reactions may be analyzed, and fragments may be purified by electrophoresis through an agarose or polyacrylamide gel, using well known methods that are routine for those skilled in the art.

GENETIC ELEMENT generally means a polynucleotide comprising a region that encodes a polypeptide or a region that regulates transcription or translation or other processes important to expression of the polypeptide in a host cell, or a polynucleotide comprising both a region that encodes a polypeptide and a region operably linked thereto that regulates expression.

Genetic elements may be comprised within a vector that replicates as an episomal element; that is, as a molecule physically independent of the host cell genome. They may be comprised within mini-chromosomes, such as those that arise during amplification of transfected DNA by methotrexate selection in eukaryotic cells. Genetic elements also may be comprised within a host cell genome; not in their natural state but, rather, following manipulation such as isolation, cloning and introduction

into a host cell in the form of purified DNA or in a vector, among others.

ISOLATED means altered "by the hand of man" from its natural state; i.e., that, if it occurs in nature, it has been changed or removed from its original environment, or both.

For example, a naturally occurring polynucleotide or a polypeptide naturally present in a living animal in its natural state is not "isolated," but the same polynucleotide or polypeptide separated from the coexisting materials of its natural state is "isolated", as the term is employed herein. For example, with respect to polynucleotides, the term isolated means that it is separated from the chromosome and cell in which it naturally occurs.

As part of or following isolation, such polynucleotides can be joined to other polynucleotides, such as DNAs, for mutagenesis, to form fusion proteins, and for propagation or expression in a host, for instance. The isolated polynucleotides, alone or joined to other polynucleotides such as vectors, can be introduced into host cells, in culture or in whole organisms. When introduced into host cells in culture or in whole organisms, such DNAs still would be isolated, as the term is used herein, because they would not be in their naturally occurring form or environment. Similarly, the polynucleotides and polypeptides may occur in a composition, such as media formulations, solutions for introduction of polynucleotides or polypeptides, for example, into cells, compositions or solutions for chemical or enzymatic reactions, for instance, which are not naturally occurring compositions, and, therein remain isolated polynucleotides or polypeptides within the meaning of that term as it is employed herein.

LIGATION refers to the process of forming

phosphodiester bonds between two or more polynucleotides, which most often are double stranded DNAS. Techniques for ligation are well known to the art and protocols for ligation are described in standard laboratory manuals and references, such as, for instance, Sambrook et al., MOLECULAR CLONING, A LABORATORY MANUAL, 2nd Ed.; Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989) and Maniatis et al., pg. 146, as cited below.

OLIGONUCLEOTIDE(S) refers to relatively short polynucleotides. Often the term refers to single-stranded deoxyribonucleotides, but it can refer as well to single-or double-stranded ribonucleotides, RNA:DNA hybrids and double-stranded DNAS, among others.

Oligonucleotides, such as single-stranded DNA probe oligonucleotides, often are synthesized by chemical methods, such as those implemented on automated oligonucleotide synthesizers. However, oligonucleotides can be made by a variety of other methods, including *in vitro* recombinant DNA-mediated techniques and by expression of DNAs in cells and organisms.

Initially, chemically synthesized DNAs typically are obtained without a 5' phosphate. The 5' ends of such oligonucleotides are not substrates for phosphodiester bond formation by ligation reactions that employ DNA ligases typically used to form recombinant DNA molecules. Where ligation of such oligonucleotides is desired, a phosphate can be added by standard techniques, such as those that employ a kinase and ATP.

The 3' end of a chemically synthesized oligonucleotide generally has a free hydroxyl group and, in the presence of a ligase, such as T4 DNA ligase, readily will form a phosphodiester bond with a 5' phosphate of another polynucleotide, such as another oligonucleotide. As is well known, this reaction can be prevented selectively, where desired, by removing the 5' phosphates of the other

polynucleotide(s) prior to ligation.

PLASMIDS generally are designated herein by a lower case "p" preceded and/or followed by capital letters and/or numbers, in accordance with standard naming conventions that are familiar to those of skill in the art. Starting plasmids disclosed herein are either commercially available, publicly available on an unrestricted basis, or can be constructed from available plasmids by routine application of well known, published procedures. Many plasmids and other cloning and expression vectors that can be used in accordance with the present invention are well known and readily available to those of skill in the art. Moreover, those of skill readily may construct any number of other plasmids suitable for use in the invention. The properties, construction and use of such plasmids, as well as other vectors, in the present invention will be readily apparent to those of skill from the present disclosure.

POLYNUCLEOTIDE(S) generally refers to any polyribonucleotide or polydeoxribonucleotide, which may be unmodified RNA or DNA or modified RNA or DNA. Thus, for instance, polynucleotides as used herein refers to, among others, single-and double-stranded DNA, DNA that is a mixture of single-and double-stranded regions, single- and double-stranded RNA, and RNA that is mixture of single- and double-stranded regions, hybrid molecules comprising DNA and RNA that may be single-stranded or, more typically, double-stranded or a mixture of single- and double-stranded regions. In addition, polynucleotide as used herein refers to triple-stranded regions comprising RNA or DNA or both RNA and DNA. The strands in such regions may be from the same molecule or from different molecules. The regions may include all of one or more of the molecules, but more typically involve only a region of some of the molecules. One of the molecules of a triple-helical region often is an oligonucleotide.

As used herein, the term polynucleotide includes DNAs or RNAs as described above that contain one or more modified bases. Thus, DNAs or RNAs with backbones modified for stability or for other reasons are "polynucleotides" as that term is intended herein. Moreover, DNAs or RNAs comprising unusual bases, such as inosine, or modified bases, such as tritylated bases, to name just two examples, are polynucleotides as the term is used herein.

It will be appreciated that a great variety of modifications have been made to DNA and RNA that serve many useful purposes known to those of skill in the art. The term polynucleotide as it is employed herein embraces such chemically, enzymatically or metabolically modified forms of polynucleotides, as well as the chemical forms of DNA and RNA characteristic of viruses and cells, including simple and complex cells, inter alia.

POLYPEPTIDES, as used herein, includes all polypeptides as described below. The basic structure of polypeptides is well known and has been described in innumerable textbooks and other publications in the art. In this context, the term is used herein to refer to any peptide or protein comprising two or more amino acids joined to each other in a linear chain by peptide bonds. As used herein, the term refers to both short chains, which also commonly are referred to in the art as peptides, oligopeptides and oligomers, for example, and to longer chains, which generally are referred to in the art as proteins, of which there are many types. it will be appreciated that polypeptides often contain amino acids other than the 20 amino acids commonly referred to as the 20 naturally occurring amino acids, and that many amino acids, including the terminal amino acids, may be modified in a given polypeptide, either by natural processes, such as processing and other post-translational modifications, but also by chemical modification techniques which are well

known to the art. Even the common modifications that occur naturally in polypeptides are too numerous to list exhaustively here, but they are well described in basic texts and in more detailed monographs, as well as in a voluminous research literature, and they are well known to those of skill in the art.

Among the known modifications which may be present in polypeptides of the present are, to name an illustrative few, acetylation, acylation, ADP-ribosylation, amidation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid derivative, covalent attachment of phosphotidylinositol, cross-linking, cyclization, disulfide bond formation, demethylation, formation of covalent cross-links, formation of cystine, formation of pyroglutamate, formylation, gamma-carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, myristoylation, oxidation, proteolytic processing, phosphorylation, prenylation, racemization, selenoylation, sulfation, transfer-RNA mediated addition of amino acids to proteins such as arginylation, and ubiquitination.

Such modifications are well known to those of skill and have been described in great detail in the scientific literature. Several particularly common modifications, glycosylation, lipid attachment, sulfation, gamma-carboxylation of glutamic acid residues, hydroxylation and ADP-ribosylation, for instance, are described in most basic texts, such as, for instance PROTEINS STRUCTURE AND MOLECULAR PROPERTIES, 2nd Ed., T. E. Creighton, W. H. Freeman and Company, New York (1993). Many detailed reviews are available on this subject, such as, for example, those provided by Wold, F., Posttranslational Protein Modifications: Perspectives and Prospects, pgs. 1-12 in POSTTRANSLATIONAL COVALENT MODIFICATION OF PROTEINS, B. C.

Johnson, Ed., Academic Press, New York (1983); Seifter et al., Analysis for protein modifications and nonprotein cofactors, Meth. Enzymol. 182: 626-646 (1990) and Rattan et al., Protein Synthesis: Posttranslational Modifications and Aging, Ann. N.Y. Acad. Sci. 663: 48-62 (1992).

It will be appreciated, as is well known and as noted above, that polypeptides are not always entirely linear. For instance, polypeptides may be branched as a result of ubiquitination, and they may be circular, with or without branching, generally as a result of posttranslation events, including natural processing event and events brought about by human manipulation which do not occur naturally. Circular, branched and branched circular polypeptides may be synthesized by non-translation natural process and by entirely synthetic methods, as well.

Modifications can occur anywhere in a polypeptide, including the peptide backbone, the amino acid side-chains and the amino or carboxyl termini. In fact, blockage of the amino or carboxyl group in a polypeptide, or both, by a covalent modification, is common in naturally occurring and synthetic polypeptides and such modifications may be present in polypeptides of the present invention, as well. For instance, the amino terminal residue of polypeptides made in *E. coli*, prior to proteolytic processing, almost invariably will be N-formylmethionine.

The modifications that occur in a polypeptide often will be a function of how it is made. For polypeptides made by expressing a cloned gene in a host, for instance, the nature and extent of the modifications in large part will be determined by the host cell posttranslational modification capacity and the modification signals present in the polypeptide amino acid sequence. For instance, as is well known, glycosylation often does not occur in bacterial hosts such as *E. coli*. Accordingly, when glycosylation is desired, a polypeptide should be expressed in a

glycosylating host, generally a eukaryotic cell. Insect cells often carry out the same posttranslational glycosylations as mammalian cells and, for this reason, insect cell expression systems have been developed to express efficiently mammalian proteins having native patterns of glycosylation, *inter alia*. Similar considerations apply to other modifications.

It will be appreciated that the same type of modification may be present in the same or varying degree at several sites in a given polypeptide. Also, a given polypeptide may contain many types of modifications.

In general, as used herein, the term polypeptide encompasses all such modifications, particularly those that are present in polypeptides synthesized by expressing a polynucleotide in a host cell.

VARIANT(S) of polynucleotides or polypeptides, as the term is used herein, are polynucleotides or polypeptides that differ from a reference polynucleotide or polypeptide, respectively. Variants in this sense are described below and elsewhere in the present disclosure in greater detail.

A polynucleotide that differs in nucleotide sequence from another, reference polynucleotide. Generally, differences are limited so that the nucleotide sequences of the reference and the variant are closely similar overall and, in many regions, identical.

As noted below, changes in the nucleotide sequence of the variant may be silent. That is, they may not alter the amino acids encoded by the polynucleotide. Where alterations are limited to silent changes of this type a variant will encode a polypeptide with the same amino acid sequence as the reference. Also as noted below, changes in the nucleotide sequence of the variant may alter the amino acid sequence of a polypeptide encoded by the reference polynucleotide. Such nucleotide changes may result in amino acid substitutions, additions, deletions, fusions and

truncations in the polypeptide encoded by the reference sequence, as discussed below.

A polypeptide that differs in amino acid sequence from another, reference polypeptide. Generally, differences are limited so that the sequences of the reference and the variant are closely similar overall and, in many region, identical.

A variant and reference polypeptide may differ in amino acid sequence by one or more substitutions, additions, deletions, fusions and truncations, which may be present in any combination.

RECEPTOR MOLECULE, as used herein, refers to molecules which bind or interact specifically with LSGS polypeptides of the present invention, preferably Lng103 and Lng104, and is inclusive not only of classic receptors, which are preferred, but also other molecules that specifically bind to or interact with polypeptides of the invention (which also may be referred to as "binding molecules" and "interaction molecules," respectively and as "LSG binding or interaction molecules", "Lng103 and Lng104 binding molecules" and "Lng103 and Lng104 interaction molecules." Binding between polypeptides of the invention and such molecules, including receptor or binding or interaction molecules may be exclusive to polypeptides of the invention, which is very highly preferred, or it may be highly specific for polypeptides of the invention, which is highly preferred, or it may be highly specific to a group of proteins that includes polypeptides of the invention, which is preferred, or it may be specific to several groups of proteins at least one of which includes polypeptides of the invention.

Receptors also may be non-naturally occurring, such as antibodies and antibody-derived reagents that bind to polypeptides of the invention.

DETAILED DESCRIPTION OF THE INVENTION

The present invention relates to novel lung specific polypeptides and polynucleotides, referred to herein as LSGs, among other things, as described in greater detail below. In particular, the invention relates to novel LSG polypeptides and polynucleotides of referred to herein as human Lng103 and Lng104, which are related by amino acid sequence homology to the rat prostatic steroid-binding protein. The invention relates especially to Lng103 and Lng104 having the nucleotide and amino acid sequences set out in SEQ ID NO:1-6 and to the Lng103 and Lng104 nucleotides and amino acid sequences of the human cDNAs in ATCC Deposit No. PTA-3032 and PTA-3033 which is herein referred to as "the deposited clones" or as the "cDNA of the deposited clones." It will be appreciated that the nucleotide and amino acid sequences set out in SEQ ID NO:1-6 were obtained by sequencing the cDNA of the deposited clone. Hence, the sequence of the deposited clone is controlling as to any discrepancies between the two and any reference to the sequences of SEQ ID NO:1, 2, 5 and 6 include reference to the sequence of the human cDNA of the deposited claim.

Polynucleotides

In accordance with one aspect of the present invention, there are provided isolated LSG polynucleotides which encode Lng103 and Lng104 polypeptides having the amino acid sequences of SEQ ID NO:3 and 4.

Using the information provided herein, such as the polynucleotide sequences set out in SEQ ID NO:1, 2, 5 and 6 a polynucleotide of the present invention encoding human Lng103 and Lng104 polypeptide may be obtained using standard cloning and screening procedures, such as those for cloning cDNAs using mRNA from cells of a human tumor as starting material.

Polynucleotides of the present invention may be in

the form of RNA, such as mRNA, or in the form of DNA, including, for instance, cDNA and genomic DNA obtained by cloning or produced by chemical synthetic techniques or by a combination thereof. The DNA may be double-stranded or single-stranded. Single-stranded DNA may be the coding strand, also known as the sense strand, or it may be the non-coding strand, also referred to as the anti-sense strand.

The coding sequence which encodes the polypeptides may be identical to the coding sequence of the polynucleotides of SEQ ID NO:1, 2, 5 or 6. It also may be a polynucleotide with a different sequence, which, as a result of the redundancy (degeneracy) of the genetic code, encodes the polypeptides of SEQ ID NO: 3 or 4.

Polynucleotides of the present invention such as SEQ ID NO: 1, 2, 5 and 6, which encode these polypeptides may include, but are not limited to the coding sequence for the mature polypeptide, by itself; the coding sequence for the mature polypeptide and additional coding sequences, such as those encoding a leader or secretory sequence, such as a pre-, or pro- or prepro-protein sequence; the coding sequence of the mature polypeptide, with or without the aforementioned additional coding sequences, together with additional, non-coding sequences, including for example, but not limited to introns and non-coding 5' and 3' sequences, such as the transcribed, non-translated sequences that play a role in transcription, mRNA processing--including splicing and polyadenylation signals, for example--ribosome binding and stability of mRNA; additional coding sequence which codes for additional amino acids, such as those which provide additional functionalities. Thus, for instance, the polypeptide may be fused to a marker sequence, such as a peptide, which facilitates purification of the fused polypeptide. In certain preferred embodiments of this aspect of the

invention, the marker sequence is a hexa-histidine peptide, such as the tag provided in the pQE vector (Qiagen, Inc., among others, many of which are commercially available. As described in Gentz et al., Proc. Natl. Acad. Sci., USA 86: 821-824 (1989), for instance, hexa-histidine provides for convenient purification of the fusion protein. The HA tag corresponds to an epitope derived of influenza hemagglutinin protein, which has been described by Wilson et al., Cell 37: 767 (1984), for instance.

In accordance with the foregoing, the term "polynucleotide encoding a polypeptide" as used herein encompasses polynucleotides which include a sequence encoding a polypeptide of the present invention, particularly human Lng103 and Lng104 having the amino acid sequences set out in SEQ ID NO: 3 and 4. The term encompasses polynucleotides that include a single continuous region or discontinuous regions encoding the polypeptide (for example, interrupted by introns) together with additional regions, that also may contain coding and/or non-coding sequences.

The present invention further relates to variants of the herein above described polynucleotides which encode for fragments, analogs and derivatives of the polypeptides having the amino acid sequences of SEQ ID NO: 3 and 4. A variant of the polynucleotide may be a naturally occurring variant such as a naturally occurring allelic variant, or it may be a variant that is not known to occur naturally. Such non-naturally occurring variants of the polynucleotide may be made by mutagenesis techniques, including those applied to polynucleotides, cells or organisms.

Among variants in this regard are variants that differ from the aforementioned polynucleotides by nucleotide substitutions, deletions or additions. The substitutions, deletions or additions may involve one or more nucleotides. The variants may be altered in coding or

non-coding regions or both. Alterations in the coding regions may produce conservative or non-conservative amino acid substitutions, deletions or additions.

Among the particularly preferred embodiments of the invention in this regard are polynucleotides encoding polypeptides having the amino acid sequence of Lng103 and Lng104 set out in SEQ ID NO: 3 and 4; variants, analogs, derivatives and fragments thereof, and fragments of the variants, analogs and derivatives.

Further particularly preferred in this regard are LSG polynucleotides encoding Lng103 and Lng104 variants, analogs, derivatives and fragments, and variants, analogs and derivatives of the fragments, which have the amino acid sequence of the Lng103 or Lng104 polypeptides of SEQ ID NO:3 and 4 in which several, a few, 5 to 10, 1 to 5, 1 to 3, 2, 1 or no amino acid residues are substituted, deleted or added, in any combination. Especially preferred among these are silent substitutions, additions and deletions, which do not alter the properties and activities of the Lng103 and Lng104. Also especially preferred in this regard are conservative substitutions. Most highly preferred are polynucleotides encoding polypeptides having the amino acid sequences of SEQ ID NO:3 and 4, without substitutions.

Further preferred embodiments of the invention are LSG polynucleotides that are at least 70% identical to a polynucleotide encoding the Lng103 and Lng104 polypeptides having the amino acid sequences set out in SEQ ID NO:3 and 4, and polynucleotides which are complementary to such polynucleotides. Alternatively, most highly preferred are LSG polynucleotides that comprise a region that is at least 80% identical to a polynucleotide encoding the Lng103 or Lng104 polypeptides of the cDNA of the deposited clone and polynucleotides complementary thereto. In this regard, LSG polynucleotides at least 90% identical to the same are

particularly preferred, and among these particularly preferred LSG polynucleotides, those with at least 95% are especially preferred. Furthermore, those with at least 97% are highly preferred among those with at least 95%, and among these those with at least 98% and at least 99% are particularly highly preferred, with at least 99% being the more preferred.

Particularly preferred embodiments in this respect, moreover, are polynucleotides which encode polypeptides which retain substantially the same biological function or activity as the mature polypeptides of SEQ ID NO:3 or 4 encoded by the human cDNA.

The present invention further relates to polynucleotides that hybridize to the herein above-described LSG sequences. In this regard, the present invention especially relates to polynucleotides which hybridize under stringent conditions to the herein above-described polynucleotides. As herein used, the term "stringent conditions" means hybridization will occur only if there is at least 95% and preferably at least 97% identity between the sequences.

As discussed additionally herein regarding polynucleotide assays of the invention, for instance, polynucleotides of the invention as discussed above, may be used as a hybridization probe for cDNA and genomic DNA to isolate full-length cDNAs and genomic clones encoding Lng103 or Lng104 and to isolate cDNA and genomic clones of other genes that have a high sequence similarity to the human Lng103 or Lng104 genes. Such probes generally will comprise at least 15 bases. Preferably, such probes will have at least 30 bases and may have at least 50 bases.

For example, the coding region of the Lng103 and Lng104 genes may be isolated by screening using the known DNA sequence to synthesize an oligonucleotide probe. A labeled oligonucleotide having a sequence complementary to

that of a gene of the present invention is then used to screen a library of human cDNA, genomic DNA or mRNA to determine which members of the library the probe hybridizes to.

The polynucleotides and polypeptides of the present invention may be employed as research reagents and materials for discovery of treatments and diagnostics to human disease, as further discussed herein relating to polynucleotide assays, *inter alia*.

The polynucleotides may encode a polypeptide which is the mature protein plus additional amino or carboxyl-terminal amino acids, or amino acids interior to the mature polypeptide (when the mature form has more than one polypeptide chain, for instance). Such sequences may play a role in processing of a protein from precursor to a mature form, may facilitate/protein trafficking, may prolong or shorten protein half-life or may facilitate manipulation of a protein for assay or production, among other things. As generally is the case *in situ*, the additional amino acids may be processed away from the mature protein by cellular enzymes.

A precursor protein, having the mature form of the polypeptide fused to one or more prosequences may be an inactive form of the polypeptide. When prosequences are removed such inactive precursors generally are activated. Some or all of the prosequences may be removed before activation. Generally, such precursors are called proproteins.

In sum, a polynucleotide of the present invention may encode a mature protein, a mature protein plus a leader sequence (which may be referred to as a preprotein), a precursor of a mature protein having one or more prosequences which are not the leader sequences of a preprotein, or a preproprotein, which is a precursor to a proprotein, having a leader sequence and one or more

prosequences, which generally are removed during processing steps that produce active and mature forms of the polypeptide.

Deposited materials

A deposit containing the LSGs human Lng103 and Lng104 cDNA has been deposited with the American Type Culture Collection, as noted above. Also as noted above, the cDNA deposit is referred to herein as "the deposited clone" or as "the cDNA of the deposited clone."

The deposited clone was deposited with the American Type Culture Collection, 10801 University Boulevard, Manassa, Va. 20110-2209, USA, on February 12, 2001 and assigned ATCC Deposit No. PTA-3032 and PTA-3033.

The deposited materials are pCMV-XL4 and pED-21d plasmids (Invitrogen, LaJolla, CA) containing the full length Lng103 and Lng104 cDNA.

The deposit has been made under the terms of the Budapest Treaty on the international recognition of the deposit of micro-organisms for purposes of patent procedure. The strain will be irrevocably and without restriction or condition released to the public upon the issuance of a patent. The deposit is provided merely as convenience to those of skill in the art and is not an admission that a deposit is required for enablement, such as that required under 35 U.S.C §112.

The sequence of the polynucleotides contained in the deposited material, as well as the amino acid sequence of the polypeptide encoded thereby, are controlling in the event of any conflict with any description of sequences herein. A license may be required to make, use or sell the deposited materials, and no such license is hereby granted.

Polypeptides

The present invention further relates to LSG polypeptides, preferably human Lng103 and Lng104 polypeptide which have the deduced amino acid sequences of

SEQ ID NO:3 and 4. The invention also relates to fragments, analogs and derivatives of these polypeptides. The terms "fragment," "derivative" and "analog" when referring to the polypeptides of SEQ ID NO: 3 and 4 means a polypeptide which retains essentially the same biological function or activity as such polypeptides. Thus, an analog includes a proprotein which can be activated by cleavage of the proprotein portion to produce an active mature polypeptide.

The polypeptide of the present invention may be a recombinant polypeptide, a natural polypeptide or a synthetic polypeptide. In certain preferred embodiments it is a recombinant polypeptide.

The fragment, derivative or analog of the polypeptide of SEQ ID NO: 3 and 4 may be (i) one in which one or more of the amino acid residues are substituted with a conserved or non-conserved amino acid residue (preferably a conserved amino acid residue) and such substituted amino acid residue may or may not be one encoded by the genetic code, or (ii) one in which one or more of the amino acid residues includes a substituent group, or (iii) one in which the mature polypeptide is fused with another compound, such as a compound to increase the half-life of the polypeptide (for example, polyethylene glycol), or (iv) one in which the additional amino acids are fused to the mature polypeptide, such as a leader or secretory sequence or a sequence which is employed for purification of the mature polypeptide or a proprotein sequence. Such fragments, derivatives and analogs are deemed to be within the scope of those skilled in the art from the teachings herein.

Among preferred variants are those that vary from a reference by conservative amino acid substitutions. Such substitutions are those that substitute a given amino acid in a polypeptide by another amino acid of like characteristics. Typically seen as conservative

substitutions are the replacements, one for another, among the aliphatic amino acids Ala, Val, Leu and Ile; interchange of the hydroxyl residues Ser and Thr, exchange of the acidic residues Asp and Glu, substitution between the amide residues Asn and Gln, exchange of the basic residues Lys and Arg and replacements among the aromatic residues Phe, Tyr.

The polypeptides and polynucleotides of the present invention are preferably provided in an isolated form, and preferably are purified to homogeneity.

The polypeptides of the present invention include the polypeptide of SEQ ID NO: 3 or 4 (in particular the mature polypeptide) as well as polypeptides which have at least 75% similarity (preferably at least 75% identity) to the polypeptide of SEQ ID NO: 3 or 4 and more preferably at least 90% similarity (more preferably at least 90% identity) to the polypeptide of SEQ ID NO:3 or 4 and still more preferably at least 95% similarity (still more preferably at least 95% identity) to the polypeptide of SEQ ID NO:3 or 4 and also include portions of such polypeptides with such portion of the polypeptide generally containing at least 30 amino acids and more preferably at least 50 amino acids.

As known in the art "similarity" between two polypeptides is determined by comparing the amino acid sequence and its conserved amino acid substitutes of one polypeptide sequence with that of a second polypeptide.

Fragments or portions of the polypeptides of the present invention may be employed for producing the corresponding full-length polypeptide by peptide synthesis; therefore, the fragments may be employed as intermediates for producing the full-length polypeptides. Fragments or portions of the polynucleotides of the present invention may be used to synthesize full-length polynucleotides of the present invention.

Fragments

Also among preferred embodiments of this aspect of the present invention are polypeptides comprising fragments of Lng103 and Lng104, most particularly fragments of the Lng103 and Lng104 having the amino acid sequence set out in SEQ ID NO: 3 and 4, and fragments of variants and derivatives of the Lng103 and Lng104 of SEQ ID NO: 3 and 4.

In this regard a fragment is a polypeptide having an amino acid sequence that entirely is the same as part but not all of the amino acid sequence of the aforementioned Lng103 and Lng104 polypeptides and variants or derivatives thereof.

Such fragments may be "free-standing," i.e., not part of or fused to other amino acids or polypeptides, or they may be comprised within a larger polypeptide of which they form a part or region. When comprised within a larger polypeptide, the presently discussed fragments most preferably form a single continuous region. However, several fragments may be comprised within a single larger polypeptide. For instance, certain preferred embodiments relate to a fragment of a Lng103 or Lng104 polypeptide of the present comprised within a precursor polypeptide designed for expression in a host and having heterologous pre and pro-polypeptide regions fused to the amino terminus of the Lng103 and Lng104 fragment and an additional region fused to the carboxyl terminus of the fragment. Therefore, fragments in one aspect of the meaning intended herein, refers to the portion or portions of a fusion polypeptide or fusion protein derived from Lng103 and Lng104.

As representative examples of polypeptide fragments of the invention, there may be mentioned those which have from about 15 to about 139 amino acids.

In this context "about" includes the particularly recited range and ranges larger or smaller by several, a few, 5, 4, 3, 2 or 1 amino acid at either extreme or at

both extremes. Highly preferred in this regard are the recited ranges plus or minus as many as 5 amino acids at either or at both extremes. Particularly highly preferred are the recited ranges plus or minus as many as 3 amino acids at either or at both the recited extremes. Especially preferred are ranges plus or minus 1 amino acid at either or at both extremes or the recited ranges with no additions or deletions. Most highly preferred of all in this regard are fragments from about 15 to about 45 amino acids.

Among especially preferred fragments of the invention are truncation mutants of Lng103 and Lng104. Truncation mutants include Lng103 and Lng104 polypeptides having the amino acid sequence of SEQ ID NO: 3 and 4, or variants or derivatives thereof, except for deletion of a continuous series of residues (that is, a continuous region, part or portion) that includes the amino terminus, or a continuous series of residues that includes the carboxyl terminus or, as in double truncation mutants, deletion of two continuous series of residues, one including the amino terminus and one including the carboxyl terminus. Fragments having the size ranges set out about also are preferred embodiments of truncation fragments, which are especially preferred among fragments generally.

Also preferred in this aspect of the invention are fragments characterized by structural or functional attributes of Lng103 and Lng104. Preferred embodiments of the invention in this regard include fragments that comprise alpha-helix and alpha-helix forming regions ("alpha-regions"), beta-sheet and beta-sheet-forming regions ("beta-regions"), turn and turn-forming regions ("turn-regions"), coil and coil-forming regions ("coil-regions"), hydrophilic regions, hydrophobic regions, alpha amphipathic regions, beta amphipathic regions, flexible regions, surface-forming regions and high antigenic index regions of Lng103 and Lng104.

Certain preferred regions in these regards are set out in SEQ ID NO: 3 and 4 and include, but are not limited to, regions of the aforementioned types identified by analysis of the amino acid sequence set out in SEQ ID NO: 3 and 4. As set out in SEQ ID NO: 3 and 4 such preferred regions include Garnier-Robson alpha-regions, beta-regions, turn-regions and coil-regions, Chou-Fasman alpha-regions, beta-regions and turn-regions, Kyte-Doolittle hydrophilic regions and hydrophilic regions, Eisenberg alpha and beta amphipathic regions, Karplus-Schulz flexible regions, Emini surface-forming regions and Jameson-Wolf high antigenic index regions.

Among highly preferred fragments in this regard are those that comprise regions of Lng103 and Lng104 that combine several structural features, such as several of the features set out above. In this regard, the regions defined by the residues of SEQ ID NO:3 and 4, which all are characterized by amino acid compositions highly characteristic of turn-regions, hydrophilic regions, flexible-regions, surface-forming regions, and high antigenic index-regions, are especially highly preferred regions. Such regions may be comprised within a larger polypeptide or may be by themselves a preferred fragment of the present invention, as discussed above. It will be appreciated that the term "about" as used in this paragraph has the meaning set out above regarding fragments in general.

Further preferred regions are those that mediate activities of Lng103 or Lng104. Most highly preferred in this regard are fragments that have a chemical, biological or other activity of Lng103 and Lng104, including those with a similar activity or an improved activity, or with a decreased undesirable activity. Highly preferred in this regard are fragments that contain regions that are homologs in sequence, or in position, or in both sequence and to

active regions of related polypeptides, such as the related polypeptides set out in SEQ ID NO:3 and 4 and which include lung specific-binding proteins. Among particularly preferred fragments in these regards are truncation mutants, as discussed above.

It will be appreciated that the invention also relates to, among others, polynucleotides encoding the aforementioned fragments, polynucleotides that hybridize to polynucleotides encoding the fragments, particularly those that hybridize under stringent conditions, and polynucleotides, such as PCR primers, for amplifying polynucleotides that encode the fragments. In these regards, preferred polynucleotides are those that correspondent to the preferred fragments, as discussed above.

Diagnostic Assays

The present invention also relates to diagnostic assays and methods, both quantitative and qualitative for detecting, diagnosing, monitoring, staging and prognosticating cancers by comparing levels of LSG in a human patient with those of LSG in a normal human control. For purposes of the present invention, what is meant by LSG levels is, among other things, native protein expressed by the gene comprising the human Lng103 and Lng104 polynucleotide sequences of SEQ ID NO: 1, 2, 5 or 6. An amino acid sequence of an LSG protein encoded by the polynucleotide sequence of SEQ ID NO:1 is depicted in SEQ ID NO:3. An amino acid sequence of an LSG protein encoded by the polynucleotide sequence of SEQ ID NO:2 is depicted in SEQ ID NO:4. By "LSG" it is also meant herein polynucleotides which, due to degeneracy in genetic coding, comprise variations in nucleotide sequence as compared to human Lng103 and Lng104 as depicted in SEQ ID NO: 1, 2, 5 or 6 but which still encode the same protein. The native protein being detected may be whole, a breakdown product, a

complex of molecules or chemically modified. In the alternative, what is meant by LSG as used herein, means the native mRNA encoded by the Lng103 or Lng104 gene comprising the polynucleotide sequence of SEQ ID NO: 1, 2, 5 or 6, levels of the Lng103 or Lng104 gene comprising the polynucleotide sequence of SEQ ID NO: 1, 2, 5 or 6 or levels of a polynucleotide which is capable of hybridizing under stringent conditions to the antisense sequence of SEQ ID NO: 1, 2, 5 or 6. Such levels are preferably determined in at least one of cells, tissues and/or bodily fluids, including determination of normal and abnormal levels. Thus, for instance, a diagnostic assay in accordance with the invention for diagnosing overexpression of LSG protein compared to normal control bodily fluids, cells, or tissue samples may be used to diagnose the presence of lung cancer.

All the methods of the present invention may optionally include determining the levels of other cancer markers as well as LSG. Other cancer markers, in addition to LSG, useful in the present invention will depend on the cancer being tested and are known to those of skill in the art.

The present invention provides methods for diagnosing the presence of lung cancer, and in particular squamous cell carcinoma, by analyzing for changes in levels of LSG in cells, tissues or bodily fluids compared with levels of LSG in cells, tissues or bodily fluids of preferably the same type from a normal human control, wherein an increase in levels of LSG in the patient versus the normal human control is associated with the presence of lung cancer.

Without limiting the instant invention, typically, for a quantitative diagnostic assay a positive result indicating the patient being tested has cancer is one in which cells, tissues or bodily fluid levels of the cancer marker, such as LSG, are at least two times higher, and

most preferably are at least five times higher, than in preferably the same cells, tissues or bodily fluid of a normal human control.

The present invention also provides a method of diagnosing metastatic lung cancer in a patient having lung cancer which has not yet metastasized for the onset of metastasis. In the method of the present invention, a human cancer patient suspected of having lung cancer which may have metastasized (but which was not previously known to have metastasized) is identified. This is accomplished by a variety of means known to those of skill in the art.

In the present invention, determining the presence of LSG levels in cells, tissues or bodily fluid, is particularly useful for discriminating between lung cancer which has not metastasized and lung cancer which has metastasized. Existing techniques have difficulty discriminating between lung cancer which has metastasized and lung cancer which has not metastasized and proper treatment selection is often dependent upon such knowledge.

In the present invention, the cancer marker levels measured in such cells, tissues or bodily fluid is LSG, and are compared with levels of LSG in preferably the same cells, tissue or bodily fluid type of a normal human control. That is, if the cancer marker being observed is just LSG in serum, this level is preferably compared with the level of LSG in serum of a normal human control. An increase in the LSG in the patient versus the normal human control is associated with lung cancer which has metastasized.

Without limiting the instant invention, typically, for a quantitative diagnostic assay a positive result indicating the cancer in the patient being tested or monitored has metastasized is one in which cells, tissues or bodily fluid levels of the cancer marker, such as LSG, are at least two times higher, and most preferably are at

least five times higher, than in preferably the same cells, tissues or bodily fluid of a normal patient.

Normal human control as used herein includes a human patient without cancer and/or non cancerous samples from the patient; in the methods for diagnosing or monitoring for metastasis, normal human control may preferably also include samples from a human patient that is determined by reliable methods to have lung cancer which has not metastasized.

Staging

The invention also provides a method of staging lung cancer in a human patient. The method comprises identifying a human patient having such cancer and analyzing cells, tissues or bodily fluid from such human patient for LSG. The LSG levels determined in the patient are then compared with levels of LSG in preferably the same cells, tissues or bodily fluid type of a normal human control, wherein an increase in LSG levels in the human patient versus the normal human control is associated with a cancer which is progressing and a decrease in the levels of LSG (but still increased over true normal levels) is associated with a cancer which is regressing or in remission.

Monitoring

Further provided is a method of monitoring lung cancer in a human patient having such cancer for the onset of metastasis. The method comprises identifying a human patient having such cancer that is not known to have metastasized; periodically analyzing cells, tissues or bodily fluid from such human patient for LSG; and comparing the LSG levels determined in the human patient with levels of LSG in preferably the same cells, tissues or bodily fluid type of a normal human control, wherein an increase in LSG levels in the human patient versus the normal human control is associated with a cancer which has metastasized.

In this method, normal human control samples may also include prior patient samples.

Further provided by this invention is a method of monitoring the change in stage of lung cancer in a human patient having such cancer. The method comprises identifying a human patient having such cancer; periodically analyzing cells, tissues or bodily fluid from such human patient for LSG; and comparing the LSG levels determined in the human patient with levels of LSG in preferably the same cells, tissues or bodily fluid type of a normal human control, wherein an increase in LSG levels in the human patient versus the normal human control is associated with a cancer which is progressing in stage and a decrease in the levels of LSG is associated with a cancer which is regressing in stage or in remission. In this method, normal human control samples may also include prior patient samples.

Monitoring a patient for onset of metastasis is periodic and preferably done on a quarterly basis. However, this may be done more or less frequently depending on the cancer, the particular patient, and the stage of the cancer.

Prognostic Testing and Clinical Trial Monitoring

The methods described herein can further be utilized as prognostic assays to identify subjects having or at risk of developing a disease or disorder associated with increased levels of LSG. The present invention provides a method in which a test sample is obtained from a human patient and LSG is detected. The presence of higher LSG levels as compared to normal human controls is diagnostic for the human patient being at risk for developing cancer, particularly lung cancer.

The effectiveness of therapeutic agents to decrease expression or activity of the LSGs of the invention can also be monitored by analyzing levels of expression of the

LSGs in a human patient in clinical trials or in *in vitro* screening assays such as in human cells. In this way, the gene expression pattern can serve as a marker, indicative of the physiological response of the human patient, or cells as the case may be, to the agent being tested.

Detection of genetic lesions or mutations

The methods of the present invention can also be used to detect genetic lesions or mutations in LSG, thereby determining if a human with the genetic lesion is at risk for lung cancer or has lung cancer. Genetic lesions can be detected, for example, by ascertaining the existence of a deletion and/or addition and/or substitution of one or more nucleotides from the LSGs of this invention, a chromosomal rearrangement of LSG, aberrant modification of LSG (such as of the methylation pattern of the genomic DNA), the presence of a non-wild type splicing pattern of a mRNA transcript of LSG, allelic loss of LSG, and/or inappropriate post-translational modification of LSG protein. Methods to detect such lesions in the LSG of this invention are known to those of skill in the art.

Assay Techniques

Assay techniques that can be used to determine levels of gene expression (including protein levels), such as LSG of the present invention, in a sample derived from a patient are well known to those of skill in the art. Such assay methods include, without limitation, radioimmunoassays, reverse transcriptase PCR (RT-PCR) assays, immunohistochemistry assays, *in situ* hybridization assays, competitive-binding assays, Western Blot analyses, ELISA assays and proteomic approaches: two-dimensional gel electrophoresis (2D electrophoresis) and non-gel based approaches such as mass spectrometry or protein interaction profiling. Among these, ELISAs are frequently preferred to diagnose a gene's expressed protein in biological fluids.

An ELISA assay initially comprises preparing an

antibody, if not readily available from a commercial source, specific to LSG, preferably a monoclonal antibody. In addition a reporter antibody generally is prepared which binds specifically to LSG. The reporter antibody is attached to a detectable reagent such as radioactive, fluorescent or enzymatic reagent, for example horseradish peroxidase enzyme or alkaline phosphatase.

To carry out the ELISA, antibody specific to LSG is incubated on a solid support, e.g. a polystyrene dish, that binds the antibody. Any free protein binding sites on the dish are then covered by incubating with a non-specific protein such as bovine serum albumin. Next, the sample to be analyzed is incubated in the dish, during which time LSG binds to the specific antibody attached to the polystyrene dish. Unbound sample is washed out with buffer. A reporter antibody specifically directed to LSG and linked to a detectable reagent such as horseradish peroxidase is placed in the dish resulting in binding of the reporter antibody to any monoclonal antibody bound to LSG. Unattached reporter antibody is then washed out. Reagents for peroxidase activity, including a colorimetric substrate are then added to the dish. Immobilized peroxidase, linked to LSG antibodies, produces a colored reaction product. The amount of color developed in a given time period is proportional to the amount of LSG protein present in the sample. Quantitative results typically are obtained by reference to a standard curve.

A competition assay can also be employed wherein antibodies specific to LSG are attached to a solid support and labeled LSG and a sample derived from the host are passed over the solid support. The amount of label detected which is attached to the solid support can be correlated to a quantity of LSG in the sample.

Using all or a portion of a nucleic acid sequence of LSG of the present invention as a hybridization probe,

nucleic acid methods can also be used to detect LSG mRNA as a marker for lung cancer. Polymerase chain reaction (PCR) and other nucleic acid methods, such as ligase chain reaction (LCR) and nucleic acid sequence based amplification (NASBA), can be used to detect malignant cells for diagnosis and monitoring of various malignancies. For example, reverse-transcriptase PCR (RT-PCR) is a powerful technique which can be used to detect the presence of a specific mRNA population in a complex mixture of thousands of other mRNA species. In RT-PCR, an mRNA species is first reverse transcribed to complementary DNA (cDNA) with use of the enzyme reverse transcriptase; the cDNA is then amplified as in a standard PCR reaction. RT-PCR can thus reveal by amplification the presence of a single species of mRNA. Accordingly, if the mRNA is highly specific for the cell that produces it, RT-PCR can be used to identify the presence of a specific type of cell.

Hybridization to clones or oligonucleotides arrayed on a solid support (i.e. gridding) can be used to both detect the expression of and quantitate the level of expression of that gene. In this approach, a cDNA encoding the LSG gene is fixed to a substrate. The substrate may be of any suitable type including but not limited to glass, nitrocellulose, nylon or plastic. At least a portion of the DNA encoding the LSG gene is attached to the substrate and then incubated with the analyte, which may be RNA or a complementary DNA (cDNA) copy of the RNA, isolated from the tissue of interest. Hybridization between the substrate bound DNA and the analyte can be detected and quantitated by several means including but not limited to radioactive labeling or fluorescence labeling of the analyte or a secondary molecule designed to detect the hybrid. Quantitation of the level of gene expression can be done by comparison of the intensity of the signal from the analyte compared with that determined from known standards. The

standards can be obtained by *in vitro* transcription of the target gene, quantitating the yield, and then using that material to generate a standard curve.

Of the proteomic approaches, 2D electrophoresis is a technique well known to those in the art. Isolation of individual proteins from a sample such as serum is accomplished using sequential separation of proteins by different characteristics usually on polyacrylamide gels. First, proteins are separated by size using an electric current. The current acts uniformly on all proteins, so smaller proteins move farther on the gel than larger proteins. The second dimension applies a current perpendicular to the first and separates proteins not on the basis of size but on the specific electric charge carried by each protein. Since no two proteins with different sequences are identical on the basis of both size and charge, the result of a 2D separation is a square gel in which each protein occupies a unique spot. Analysis of the spots with chemical or antibody probes, or subsequent protein microsequencing can reveal the relative abundance of a given protein and the identity of the proteins in the sample.

The above tests can be carried out on samples derived from a variety of cells, bodily fluids and/or tissue extracts such as homogenates or solubilized tissue obtained from a patient. Tissue extracts are obtained routinely from tissue biopsy and autopsy material. Bodily fluids useful in the present invention include blood, urine, saliva or any other bodily secretion or derivative thereof. By blood it is meant to include whole blood, plasma, serum or any derivative of blood.

In Vivo Targeting of LSG/Lung Cancer Therapy

Identification of this LSG is also useful in the rational design of new therapeutics for imaging and treating cancers, and in particular lung cancer. For

example, in one embodiment, antibodies which specifically bind to LSG can be raised and used *in vivo* in patients suspected of suffering from lung cancer. Antibodies which specifically bind LSG can be injected into a patient suspected of having lung cancer for diagnostic and/or therapeutic purposes. Thus, another aspect of the present invention provides for a method for preventing the onset and treatment of lung cancer in a human patient in need of such treatment by administering to the patient an effective amount of antibody. By "effective amount" it is meant the amount or concentration of antibody needed to bind to the target antigens expressed on the tumor to cause tumor shrinkage for surgical removal, or disappearance of the tumor. The binding of the antibody to the overexpressed LSG is believed to cause the death of the cancer cell expressing such LSG. The preparation and use of antibodies for *in vivo* diagnosis and treatment is well known in the art. For example, antibody-chelators labeled with Indium-111 have been described for use in the radioimmunoscintographic imaging of carcinoembryonic antigen expressing tumors (Sumerdon et al. Nucl. Med. Biol. 1990 17:247-254). In particular, these antibody-chelators have been used in detecting tumors in patients suspected of having recurrent colorectal cancer (Griffin et al. J. Clin. Onc. 1991 9:631-640). Antibodies with paramagnetic ions as labels for use in magnetic resonance imaging have also been described (Lauffer, R.B. Magnetic Resonance in Medicine 1991 22:339-342). Antibodies directed against LSG can be used in a similar manner. Labeled antibodies which specifically bind LSG can be injected into patients suspected of having lung cancer for the purpose of diagnosing or staging of the disease status of the patient. The label used will be selected in accordance with the imaging modality to be used. For example, radioactive labels such as Indium-111, Technetium-99m or Iodine-131 can

be used for planar scans or single photon emission computed tomography (SPECT). Positron emitting labels such as Fluorine-19 can be used in positron emission tomography. Paramagnetic ions such as Gadlinium (III) or Manganese (II) can be used in magnetic resonance imaging (MRI). Presence of the label, as compared to imaging of normal tissue, permits determination of the spread of the cancer. The amount of label within an organ or tissue also allows determination of the presence or absence of cancer in that organ or tissue.

Antibodies which can be used in *in vivo* methods include polyclonal, monoclonal and omniclonal antibodies and antibodies prepared via molecular biology techniques. Antibody fragments and aptamers and single-stranded oligonucleotides such as those derived from an *in vitro* evolution protocol referred to as SELEX and well known to those skilled in the art can also be used.

Screening Assays

The present invention also provides methods for identifying modulators which bind to LSG protein or have a modulatory effect on the expression or activity of LSG protein. Modulators which decrease the expression or activity of LSG protein are believed to be useful in treating lung cancer. Such screening assays are known to those of skill in the art and include, without limitation, cell-based assays and cell free assays.

Small molecules predicted via computer imaging to specifically bind to regions of LSG can also be designed, synthesized and tested for use in the imaging and treatment of lung cancer. Further, libraries of molecules can be screened for potential anticancer agents by assessing the ability of the molecule to bind to the LSGs identified herein. Molecules identified in the library as being capable of binding to LSG are key candidates for further evaluation for use in the treatment of lung cancer. In a

preferred embodiment, these molecules will downregulate expression and/or activity of LSG in cells.

Adoptive Immunotherapy and Vaccines

Adoptive immunotherapy of cancer refers to a therapeutic approach in which immune cells with an antitumor reactivity are administered to a tumor-bearing host, with the aim that the cells mediate either directly or indirectly, the regression of an established tumor. Transfusion of lymphocytes, particularly T lymphocytes, falls into this category and investigators at the National Cancer Institute (NCI) have used autologous reinfusion of peripheral blood lymphocytes or tumor-infiltrating lymphocytes (TIL), T cell cultures from biopsies of subcutaneous lymph nodules, to treat several human cancers (Rosenberg, S. A., U.S. Patent No. 4,690,914, issued Sep. 1, 1987; Rosenberg, S. A., et al., 1988, N. England J. Med. 319:1676-1680).

The present invention relates to compositions and methods of adoptive immunotherapy for the prevention and/or treatment of primary and metastatic lung cancer in humans using macrophages sensitized to the antigenic LSG molecules, with or without non-covalent complexes of heat shock protein (hsp). Antigenicity or immunogenicity of the LSG is readily confirmed by the ability of the LSG protein or a fragment thereof to raise antibodies or educate naive effector cells, which in turn lyse target cells expressing the antigen (or epitope).

Cancer cells are, by definition, abnormal and contain proteins which should be recognized by the immune system as foreign since they are not present in normal tissues. However, the immune system often seems to ignore this abnormality and fails to attack tumors. The foreign LSG proteins that are produced by the cancer cells can be used to reveal their presence. The LSG is broken into short fragments, called tumor antigens, which are displayed on

the surface of the cell. These tumor antigens are held or presented on the cell surface by molecules called MHC, of which there are two types: class I and II. Tumor antigens in association with MHC class I molecules are recognized by cytotoxic T cells while antigen-MHC class II complexes are recognized by a second subset of T cells called helper cells. These cells secrete cytokines which slow or stop tumor growth and help another type of white blood cell, B cells, to make antibodies against the tumor cells.

In adoptive immunotherapy, T cells or other antigen presenting cells (APCs) are stimulated outside the body (*ex vivo*), using the tumor specific LSG antigen. The stimulated cells are then reinfused into the patient where they attack the cancerous cells. Research has shown that using both cytotoxic and helper T cells is far more effective than using either subset alone. Additionally, the LSG antigen may be complexed with heat shock proteins to stimulate the APCs as described in U.S. Patent No. 5,985,270.

The APCs can be selected from among those antigen presenting cells known in the art including, but not limited to, macrophages, dendritic cells, B lymphocytes, and a combination thereof, and are preferably macrophages. In a preferred use, wherein cells are autologous to the individual, autologous immune cells such as lymphocytes, macrophages or other APCs are used to circumvent the issue of whom to select as the donor of the immune cells for adoptive transfer. Another problem circumvented by use of autologous immune cells is graft versus host disease which can be fatal if unsuccessfully treated.

In adoptive immunotherapy with gene therapy, DNA of the LSG can be introduced into effector cells similarly as in conventional gene therapy. This can enhance the cytotoxicity of the effector cells to tumor cells as they have been manipulated to produce the antigenic protein

resulting in improvement of the adoptive immunotherapy.

LSG antigens of this invention are also useful as components of lung cancer vaccines. The vaccine comprises an immunogenically stimulatory amount of an LSG antigen. Immunogenically stimulatory amount refers to that amount of antigen that is able to invoke the desired immune response in the recipient for the amelioration, or treatment of lung cancer. Effective amounts may be determined empirically by standard procedures well known to those skilled in the art.

The LSG antigen may be provided in any one of a number of vaccine formulations which are designed to induce the desired type of immune response, e.g., antibody and/or cell mediated. Such formulations are known in the art and include, but are not limited to, formulations such as those described in U.S. Patent 5,585,103. Vaccine formulations of the present invention used to stimulate immune responses can also include pharmaceutically acceptable adjuvants.

Vectors, host cells, expression

The present invention also relates to vectors which include polynucleotides of the present invention, host cells which are genetically engineered with vectors of the invention and the production of polypeptides of the invention by recombinant techniques.

Host cells can be genetically engineered to incorporate polynucleotides and express polypeptides of the present invention. For instance, polynucleotides may be introduced into host cells using well known techniques of infection, transduction, transfection, transvection and transformation. The polynucleotides may be introduced alone or with other polynucleotides. Such other polynucleotides may be introduced independently, co-introduced or introduced joined to the polynucleotides of the invention.

Thus, for instance, polynucleotides of the invention may be transfected into host cells with another, separate, polynucleotide encoding a selectable marker, using standard

techniques for co-transfection and selection in, for instance, mammalian cells. In this case the polynucleotides generally will be stably incorporated into the host cell genome.

Alternatively, the polynucleotides may be joined to a vector containing a selectable marker for propagation in a host. The vector construct may be introduced into host cells by the aforementioned techniques. Generally, a plasmid vector is introduced as DNA in a precipitate, such as a calcium phosphate precipitate, or in a complex with a charged lipid. Electroporation also may be used to introduce polynucleotides into a host. If the vector is a virus, it may be packaged *in vitro* or introduced into a packaging cell and the packaged virus may be transduced into cells. A wide variety of techniques suitable for making polynucleotides and for introducing polynucleotides into cells in accordance with this aspect of the invention are well known and routine to those of skill in the art Such techniques are reviewed at length in Sambrook et al. cited above, which is illustrative of the many laboratory manuals that detail these techniques. In accordance with this aspect of the invention the vector may be, for example, a plasmid vector, a single- or double-stranded phage vector, a single- or double-stranded RNA or DNA viral vector. Such vectors may be introduced into cells as polynucleotides, preferably DNA, by well known techniques for introducing DNA and RNA into cells. The vectors, in the case of phage and viral vectors also may be and preferably are introduced into cells as packaged or encapsidated virus by well known techniques for infection and transduction. Viral vectors may be replication competent or replication defective. In the latter case viral propagation generally will occur only in complementing host cells.

Preferred among vectors, in certain respects, are those for expression of polynucleotides and polypeptides of

the present invention. Generally, such vectors comprise cis-acting control regions effective for expression in a host operatively linked to the polynucleotide to be expressed. Appropriate trans-acting factors either are supplied by the host, supplied by a complementing vector or supplied by the vector itself upon introduction into the host.

In certain preferred embodiments in this regard, the vectors provide for specific expression. Such specific expression may be inducible expression or expression only in certain types of cells or both inducible and cell-specific. Particularly preferred among inducible vectors are vectors that can be induced for expression by environmental factors that are easy to manipulate, such as temperature and nutrient additives. A variety of vectors suitable to this aspect of the invention, including constitutive and inducible expression vectors for use in prokaryotic and eukaryotic hosts, are well known and employed routinely by those of skill in the art.

The engineered host cells can be cultured in conventional nutrient media, which may be modified as appropriate for, inter alia, activating promoters, selecting transformants or amplifying genes. Culture conditions, such as temperature, pH and the like, previously used with the host cell selected for expression generally will be suitable for expression of polypeptides of the present invention as will be apparent to those of skill in the art.

A great variety of expression vectors can be used to express a polypeptide of the invention. Such vectors include chromosomal, episomal and virus-derived vectors e.g., vectors derived from bacterial plasmids, from bacteriophage, from yeast episomes, from yeast chromosomal elements, from viruses such as baculoviruses, papova viruses, such as SV40, vaccinia viruses, adenoviruses, fowl

pox viruses, pseudorabies viruses and retroviruses, and vectors derived from combinations thereof, such as those derived from plasmid and bacteriophage genetic elements, such as cosmids and phagemids, all may be used for expression in accordance with this aspect of the present invention. Generally, any vector suitable to maintain, propagate or express polynucleotides to express a polypeptide in a host may be used for expression in this regard.

The appropriate DNA sequence may be inserted into the vector by any of a variety of well-known and routine techniques. In general, a DNA sequence for expression is joined to an expression vector by cleaving the DNA sequence and the expression vector with one or more restriction endonucleases and then joining the restriction fragments together using T4 DNA ligase. Procedures for restriction and ligation that can be used to this end are well known and routine to those of skill. Suitable procedures in this regard, and for constructing expression vectors using alternative techniques, which also are well known and routine to those skill, are set forth in great detail in Sambrook et al. cited elsewhere herein.

The DNA sequence in the expression vector is operatively linked to appropriate expression control sequence(s), including, for instance, a promoter to direct mRNA transcription. Representatives of such promoters include the phage lambda PL promoter, the *E. coli* lac, trp and tac promoters, the SV40 early and late promoters and promoters of retroviral LTRs, to name just a few of the well-known promoters. It will be understood that numerous promoters not mentioned are suitable for use in this aspect of the invention are well known and readily may be employed by those of skill in the manner illustrated by the discussion and the examples herein.

In general, expression constructs will contain sites

for transcription initiation and termination, and, in the transcribed region, a ribosome binding site for translation. The coding portion of the mature transcripts expressed by the constructs will include a translation initiating AUG at the beginning and a termination codon appropriately positioned at the end of the polypeptide to be translated.

In addition, the constructs may contain control regions that regulate as well as engender expression. Generally, in accordance with many commonly practiced procedures, such regions will operate by controlling transcription, such as repressor binding sites and enhancers, among others.

Vectors for propagation and expression generally will include selectable markers. Such markers also may be suitable for amplification or the vectors may contain additional markers for this purpose. In this regard, the expression vectors preferably contain one or more selectable marker genes to provide a phenotypic trait for selection of transformed host cells. Preferred markers include dihydrofolate reductase or neomycin resistance for eukaryotic cell culture, and tetracycline or ampicillin resistance genes for culturing *E. coli* and other bacteria.

The vector containing the appropriate DNA sequence as described elsewhere herein, as well as an appropriate promoter, and other appropriate control sequences, may be introduced into an appropriate host using a variety of well known techniques suitable to expression therein of a desired polypeptide. Representative examples of appropriate hosts include bacterial cells, such as *E. coli*, *Streptomyces* and *Salmonella typhimurium* cells; fungal cells, such as yeast cells; insect cells such as *Drosophila* S2 and *Spodoptera* Sf9 cells; animal cells such as CHO, COS and Bowes melanoma cells; and plant cells. Hosts for a great variety of expression constructs are well known, and

those of skill will be enabled by the present disclosure readily to select a host for expressing a polypeptides in accordance with this aspect of the present invention.

More particularly, the present invention also includes recombinant constructs, such as expression constructs, comprising one or more of the sequences described above. The constructs comprise a vector, such as a plasmid or viral vector, into which such a sequence of the invention has been inserted. The sequence may be inserted in a forward or reverse orientation. In certain preferred embodiments in this regard, the construct further comprises regulatory sequences, including, for example, a promoter, operably linked to the sequence. Large numbers of suitable vectors and promoters are known to those of skill in the art, and there are many commercially available vectors suitable for use in the present invention.

The following vectors, which are commercially available, are provided by way of example. Among vectors preferred for use in bacteria are pQE70, pQE60 and pQE-9, available from Qiagen; pBS vectors, Phagescript vectors, Bluescript vectors, pNH8A, pNH16a, pNH18A, pNH46A, available from Stratagene; and ptrc99a, pKK223-3, pKK233-3, pDR540, pRIT5 available from Pharmacia. Among preferred eukaryotic vectors are PWLNEO, pSV2CAT, pOG44, pXT1 and pSG available from Stratagene; and pSVK3, pBPV, pMSG and pSVL available from Pharmacia. These vectors are listed solely by way of illustration of the many commercially available and well known vectors that are available to those of skill in the art for use in accordance with this aspect of the present invention. It will be appreciated that any other plasmid or vector suitable for, for example, introduction, maintenance, propagation or expression of a polynucleotide or polypeptide of the invention in a host may be used in this aspect of the invention.

Promoter regions can be selected from any desired

gene using vectors that contain a reporter transcription unit lacking a promoter region, such as a chloramphenicol acetyl transferase ("cat") transcription unit, downstream of restriction site or sites for introducing a candidate promoter fragment; i.e., a fragment that may contain a promoter. As is well known, introduction into the vector of a promoter-containing fragment at the restriction site upstream of the cat gene engenders production of CAT activity, which can be detected by standard CAT assays. Vectors suitable to this end are well known and readily available. Two such vectors are pKK232-8 and pCM7. Thus, promoters for expression of polynucleotides of the present invention include not only well known and readily available promoters, but also promoters that readily may be obtained by the foregoing technique, using a reporter gene.

Among known bacterial promoters suitable for expression of polynucleotides and polypeptides in accordance with the present invention are the *E. coli* laci and lacZ and promoters, the T3 and T7 promoters, the gpt promoter, the lambda PR, PL promoters and the trp promoter. Among known eukaryotic promoters suitable in this regard are the CMV immediate early promoter, the HSV thymidine kinase promoter, the early and late SV40 promoters, the promoters of retroviral LTRs, such as those of the Rous sarcoma virus ("RSV"), and metallothionein promoters, such as the mouse metallothionein-I promoter.

Selection of appropriate vectors and promoters for expression in a host cell is a well known procedure and the requisite techniques for expression vector construction, introduction of the vector into the host and expression in the host are routine skills in the art.

The present invention also relates to host cells containing the above-described constructs discussed above. The host cell can be a higher eukaryotic cell, such as a mammalian cell, or a lower eukaryotic cell, such as a yeast

cell, or the host cell can be a prokaryotic cell, such as a bacterial cell.

Introduction of the construct into the host cell can be effected by calcium phosphate transfection, DEAE-dextran mediated transfection, cationic lipid-mediated transfection, electroporation, transduction, infection or other methods. Such methods are described in many standard laboratory manuals, such as Davis et al. BASIC METHODS IN MOLECULAR BIOLOGY, (1986).

Constructs in host cells can be used in a conventional manner to produce the gene product encoded by the recombinant sequence. Alternatively, the polypeptides of the invention can be synthetically produced by conventional peptide synthesizers.

Mature proteins can be expressed in mammalian cells, yeast, bacteria, or other cells under the control of appropriate promoters. Cell-free translation systems can also be employed to produce such proteins using RNAs derived from the DNA constructs of the present invention. Appropriate cloning and expression vectors for use with prokaryotic and eukaryotic hosts are described by Sambrook et al., MOLECULAR CLONING: A LABORATORY MANUAL, 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989).

Generally, recombinant expression vectors will include origins of replication, a promoter derived from a highly-expressed gene to direct transcription of a downstream structural sequence, and a selectable marker to permit isolation of vector containing cells after exposure to the vector. Among suitable promoters are those derived from the genes that encode glycolytic enzymes such as 3-phosphoglycerate kinase ("PGK"), a-factor, acid phosphatase, and heat shock proteins, among others. Selectable markers include the ampicillin resistance gene of *E. coli* and the *trpl* gene of *S. cerevisiae*.

Transcription of the DNA encoding the polypeptides of the present invention by higher eukaryotes may be increased by inserting an enhancer sequence into the vector. Enhancers are cis-acting elements of DNA, usually about from 10 to 300 bp that act to increase transcriptional activity of a promoter in a given host cell-type. Examples of enhancers include the SV40 enhancer, which is located on the late side of the replication origin at bp 100 to 270, the cytomegalovirus early promoter enhancer, the polyoma enhancer on the late side of the replication origin, and adenovirus enhancers.

Polynucleotides of the invention, encoding the heterologous structural sequence of a polypeptide of the invention generally will be inserted into the vector using standard techniques so that it is operably linked to the promoter for expression. The polynucleotide will be positioned so that the transcription start site is located appropriately 5' to a ribosome binding site. The ribosome binding site will be 5' to the AUG that initiates translation of the polypeptide to be expressed. Generally, there will be no other open reading frames that begin with an initiation codon, usually AUG, and lie between the ribosome binding site and the initiating AUG. Also, generally, there will be a translation stop codon at the end of the polypeptide and there will be a polyadenylation signal and a transcription termination signal appropriately disposed at the 3' end of the transcribed region.

For secretion of the translated protein into the lumen of the endoplasmic reticulum, into the periplasmic space or into the extracellular environment, appropriate secretion signals may be incorporated into the expressed polypeptide. The signals may be endogenous to the polypeptide or they may be heterologous signals.

The polypeptide may be expressed in a modified form, such as a fusion protein, and may include not only

secretion signals but also additional heterologous functional regions. Thus, for instance, a region of additional amino acids, particularly charged amino acids, may be added to the N-terminus of the polypeptide to improve stability and persistence in the host cell, during purification or during subsequent handling and storage. Also, region also may be added to the polypeptide to facilitate purification. Such regions may be removed prior to final preparation of the polypeptide. The addition of peptide moieties to polypeptides to engender secretion or excretion, to improve stability and to facilitate purification, among others, are familiar and routine techniques in the art.

Suitable prokaryotic hosts for propagation, maintenance or expression of polynucleotides and polypeptides in accordance with the invention include *Escherichia coli*, *Bacillus subtilis* and *Salmonella typhimurium*. Various species of *Pseudomonas*, *Streptomyces*, and *Staphylococcus* are suitable hosts in this regard. Moreover, many other hosts also known to those of skill may be employed in this regard.

As a representative but non-limiting example, useful expression vectors for bacterial use can comprise a selectable marker and bacterial origin of replication derived from commercially available plasmids comprising genetic elements of the well known cloning vector pBR322. Such commercial vectors include, for example, pKK223-3 (Pharmacia Fine Chemicals, Uppsala, Sweden) and GEM1 (Promega Biotec, Madison, Wis., USA). These pBR322 "backbone" sections are combined with an appropriate promoter and the structural sequence to be expressed.

Following transformation of a suitable host strain and growth of the host strain to an appropriate cell density, where the selected promoter is inducible it is induced by appropriate means (e.g., temperature shift or

exposure to chemical inducer) and cells are cultured for an additional period.

Cells typically then are harvested by centrifugation, disrupted by physical or chemical means, and the resulting crude extract retained for further purification. Microbial cells employed in expression of proteins can be disrupted by any convenient method, including freeze-thaw cycling, sonication, mechanical disruption, or use of cell lysing agents, such methods are well know to those skilled in the art.

Various mammalian cell culture systems can be employed for expression, as well. Examples of mammalian expression systems include the COS-7 lines of monkey kidney fibroblast, described in Gluzman et al., Cell 23: 175 (1981). Other cell lines capable of expressing a compatible vector include for example, the C127, 3T3, CHO, HeLa, human kidney 293 and BHK cell lines.

Mammalian expression vectors will comprise an origin of replication, a suitable promoter and enhancer, and also any necessary ribosome binding sites, polyadenylation sites, splice donor and acceptor sites, transcriptional termination sequences, and 5' flanking non-transcribed sequences that are necessary for expression. In certain preferred embodiments in this regard DNA sequences derived from the SV40 splice sites, and the SV40 polyadenylation sites are used for required non-transcribed genetic elements of these types.

The Lng103 and Lng104 polypeptide can be recovered and purified from recombinant cell cultures by well-known methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Most preferably, high performance liquid

chromatography ("HPLC") is employed for purification. Well known techniques for refolding protein may be employed to regenerate active conformation when the polypeptide is denatured during isolation and or purification.

Polypeptides of the present invention include naturally purified products, products of chemical synthetic procedures, and products produced by recombinant techniques from a prokaryotic or eukaryotic host, including, for example, bacterial, yeast, higher plant, insect and mammalian cells. Depending upon the host employed in a recombinant production procedure, the polypeptides of the present invention may be glycosylated or may be non-glycosylated. In addition, polypeptides of the invention may also include an initial modified methionine residue, in some cases as a result of host-mediated processes.

Lng103 and Lng104 polynucleotides and polypeptides may be used in accordance with the present invention for a variety of applications, particularly those that make use of the chemical and biological properties Lng103 and Lng104. Additional applications relate to diagnosis and to treatment of disorders of cells, tissues and organisms. These aspects of the invention are illustrated further by the following discussion.

Polynucleotide assays

This invention is also related to the use of the LSG Lng103 and Lng104 polynucleotides to detect complementary polynucleotides such as, for example, as a diagnostic reagent. Detection of a mutated form of Lng103 and Lng104 associated with a dysfunction will provide a diagnostic tool that can add or define a diagnosis of a disease or susceptibility to a disease which results from under-expression, over-expression or altered expression of Lng103 and Lng104, such as, for example, a susceptibility to inherited lung cancer.

Individuals carrying mutations in the human Lng103

and Lng104 gene may be detected at the DNA level by a variety of techniques. Nucleic acids for diagnosis may be obtained from a patient's cells, such as from blood, urine, saliva, tissue biopsy and autopsy material. The genomic DNA may be used directly for detection or may be amplified enzymatically by using PCR prior to analysis. PCR (Saiki et al., Nature, 324: 163-166 (1986)). RNA or cDNA may also be used in the same ways. As an example, PCR primers complementary to the nucleic acid encoding Lng103 and Lng104 can be used to identify and analyze Lng103 and Lng104 expression and mutations. For example, deletions and insertions can be detected by a change in size of the amplified product in comparison to the normal genotype. Point mutations can be identified by hybridizing amplified DNA to radiolabeled Lng103 and Lng104 RNA or alternatively, radiolabeled Lng103 and Lng104 antisense DNA sequences. Perfectly matched sequences can be distinguished from mismatched duplexes by RNase A digestion or by differences in melting temperatures.

Sequence differences between a reference gene and genes having mutations also may be revealed by direct DNA sequencing. In addition, cloned DNA segments may be employed as probes to detect specific DNA segments. The sensitivity of such methods can be greatly enhanced by appropriate use of PCR or another amplification method. For example, a sequencing primer is used with double-stranded PCR product or a single-stranded template molecule generated by a modified PCR. The sequence determination is performed by conventional procedures with radiolabeled nucleotide or by automatic sequencing procedures with fluorescent-tags.

Genetic testing based on DNA sequence differences may be achieved by detection of alteration in electrophoretic mobility of DNA fragments in gels, with or without denaturing agents. Small sequence deletions and insertions

can be visualized by high resolution gel electrophoresis. DNA fragments of different sequences may be distinguished on denaturing formamide gradient gels in which the mobilities of different DNA fragments are retarded in the gel at different positions according to their specific melting or partial melting temperatures (see, e.g., Myers et al., Science, 230: 1242 (1985)).

Sequence changes at specific locations also may be revealed by nuclease protection assays, such as RNase and S1 protection or the chemical cleavage method (e.g., Cotton et al., Proc. Natl. Acad. Sci., USA, 85: 4397-4401 (1985)).

Thus, the detection of a specific DNA sequence may be achieved by methods such as hybridization, RNase protection, chemical cleavage, direct DNA sequencing or the use of restriction enzymes, (e.g., restriction fragment length polymorphisms ("RFLP") and Southern blotting of genomic DNA.

In addition to more conventional gel-electrophoresis and DNA sequencing, mutations also can be detected by in situ analysis.

Chromosome assays

The sequences of the present invention are also valuable for chromosome identification. The sequence is specifically targeted to and can hybridize with a particular location on an individual human chromosome. Moreover, there is a current need for identifying particular sites on the chromosome. Few chromosome marking reagents based on actual sequence data (repeat polymorphisms) are presently available for marking chromosomal location. The mapping of DNAs to chromosomes according to the present invention is an important first step in correlating those sequences with genes associated with disease.

In certain preferred embodiments in this regard, the cDNA herein disclosed is used to clone genomic DNA of a

Lng103 or Lng104 gene. This can be accomplished using a variety of well known techniques and libraries, which generally are available commercially. The genomic DNA the is used for in situ chromosome mapping using well known techniques for this purpose. Typically, in accordance with routine procedures for chromosome mapping, some trial and error may be necessary to identify a genomic probe that gives a good in situ hybridization signal.

In some cases, in addition, sequences can be mapped to chromosomes by preparing PCR primers (preferably 15-25 bp) from the cDNA. Computer analysis of the 3' untranslated region of the gene is used to rapidly select primers that do not span more than one exon in the genomic DNA, thus complicating the amplification process. These primers are then used for PCR screening of somatic cell hybrids containing individual human chromosomes. Only those hybrids containing the human gene corresponding to the primer will yield an amplified fragment.

PCR mapping of somatic cell hybrids is a rapid procedure for assigning a particular DNA to a particular chromosome. Using the present invention with the same oligonucleotide primers, sublocalization can be achieved with panels of fragments from specific chromosomes or pools of large genomic clones in an analogous manner. Other mapping strategies that can similarly be used to map to its chromosome include in situ hybridization, prescreening with labeled flow-sorted chromosomes and preselection by hybridization to construct chromosome specific-cDNA libraries.

Fluorescence *in situ* hybridization ("FISH") of a cDNA clone to a metaphase chromosomal spread can be used to provide a precise chromosomal location in one step. This technique can be used with cDNA as short as 50 or 60. For a review of this technique, see Verma et al., HUMAN CHROMOSOMES: A MANUAL OF BASIC TECHNIQUES, Pergamon Press,

New York (1988).

Once a sequence has been mapped to a precise chromosomal location, the physical position of the sequence on the chromosome can be correlated with genetic map data. Such data are found, for example, in V. McKusick, MENDELIAN INHERITANCE IN MAN, available on line through Johns Hopkins University, Welch Medical Library. The relationship between genes and diseases that have been mapped to the same chromosomal region are then identified through linkage analysis (coinheritance of physically adjacent genes).

Next, it is necessary to determine the differences in the cDNA or genomic sequence between affected and unaffected individuals. If a mutation is observed in some or all of the affected individuals but not in any normal individuals, then the mutation is likely to be the causative agent of the disease.

With current resolution of physical mapping and genetic mapping techniques, a cDNA precisely localized to a chromosomal region associated with the disease could be one of between 50 and 500 potential causative genes. (This assumes 1 megabase mapping resolution and one gene per 20 kb).

Polypeptide assays

The present invention also relates to a diagnostic assays such as quantitative and diagnostic assays for detecting levels of Lng103 and Lng104 protein in cells and tissues, and biological fluids such, for example, as blood and urine, including determination of normal and abnormal levels. Thus, for instance, a diagnostic assay in accordance with the invention for detecting over-expression or under-expression of Lng103 and Lng104 protein compared to normal control tissue samples may be used to detect the presence of neoplasia, for example. Assay techniques that can be used to determine levels of a protein, such as an Lng103 and Lng104 protein of the present invention, in a

sample derived from a host are well-known to those of skill in the art. Such assay methods include radioimmunoassays, competitive-binding assays, Western Blot analysis and ELISA assays. Among these ELISAs frequently are preferred. An ELISA assay initially comprises preparing an antibody specific to Lng103 or Lng104, preferably a monoclonal antibody. In addition a reporter antibody generally is prepared which binds to the monoclonal antibody. The reporter antibody is attached a detectable reagent such as radioactive, fluorescent or enzymatic reagent, for example horseradish peroxidase enzyme.

To carry out an ELISA a sample is removed from a host and incubated on a solid support, e.g. a polystyrene dish, that binds the proteins in the sample. Any free protein binding sites on the dish are then covered by incubating with a non-specific protein such as bovine serum albumin. Next, the monoclonal antibody is incubated in the dish during which time the monoclonal antibodies attach to any Lng103 or Lng104 proteins attached to the polystyrene dish. Unbound monoclonal antibody is washed out with buffer. The reporter antibody linked to horseradish peroxidase is placed in the dish resulting in binding of the reporter antibody to any monoclonal antibody bound to Lng103 or Lng104. Unattached reporter antibody is then washed out. Reagents for peroxidase activity, including a calorimetric substrate are then added to the dish. Immobilized peroxidase, linked to Lng103 or Lng104 through the primary and secondary antibodies, produces a colored reaction product. The amount of color developed in a given time period indicates the amount of Lng103 or Lng104 protein present in the sample. Quantitative results typically are obtained by reference to a standard curve.

A competition assay may be employed wherein antibodies specific to Lng103 or Lng104 attached to a solid support and labeled Lng103 or Lng104 and a sample derived

from the host are passed over the solid support and the amount of label detected attached to the solid support can be correlated to a quantity of Lng103 or Lng104 in the sample.

Antibodies

The polypeptides, their fragments or other derivatives, or analogs thereof, or cells expressing them can be used as an immunogen to produce antibodies thereto. These antibodies can be, for example, polyclonal or monoclonal antibodies. The present invention also includes chimeric, single chain, and humanized antibodies, as well as Fab fragments, or the product of an Fab expression library. Various procedures known in the art may be used for the production of such antibodies and fragments.

Antibodies generated against the polypeptides corresponding to a sequence of the present invention can be obtained by direct injection of the polypeptides into an animal or by administering the polypeptides to an animal, preferably a nonhuman. The antibody so obtained will then bind the polypeptides itself. In this manner, even a sequence encoding only a fragment of the polypeptides can be used to generate antibodies binding the whole native polypeptides. Such antibodies can then be used to isolate the polypeptide from tissue expressing that polypeptide.

For preparation of monoclonal antibodies, techniques which provides antibodies produced by continuous cell line cultures can be used. Examples include the hybridoma technique (Kohler, G. and Milstein, C., Nature 256: 495-497 (1975), the trioma technique, the human B-cell hybridoma technique (Kozbor et al., Immunology Today 4: 72 (1983) and the EBV-hybridoma technique to produce human monoclonal antibodies (Cole et al., pg. 77-96 in MONOCLONAL ANTIBODIES AND CANCER THERAPY, Alan R. Liss, Inc. (1985).

Techniques described for the production of single chain antibodies (U.S. Pat. No. 4,946,778) can be adapted

to produce single chain antibodies to immunogenic polypeptide products of this invention. Also, transgenic mice, or other organisms such as other mammals, may be used to express humanized antibodies to immunogenic polypeptide products of this invention.

The above-described antibodies may be employed to isolate or to identify clones expressing the polypeptide or purify the polypeptide of the present invention by attachment of the antibody to a solid support for isolation and/or purification by affinity chromatography.

Thus, among others, the polynucleotides and polypeptides of the present invention may be employed to prevent and/or treat inflammation, asthma, rhinitis, cystic fibrosis, airway disease, prevent and/or treat neoplasia, atopy, inhibit phospholipase A, sub.2, bind polychlorated biphenyls, reduce foreign protein antigenicity, inhibit monocyte and neutrophil chemotaxis and phagocytosis, inhibit platelet aggregation, regulate eicosanoid levels in the human uterus, control the growth of endometrial cells.

Lng103 and Lng104 binding molecules and assays

This invention also provides a method for identification of molecules, such as receptor molecules, that bind Lng103 and Lng104. Genes encoding proteins that bind Lng103 and Lng104, such as receptor proteins, can be identified by numerous methods known to those of skill in the art, for example, ligand panning and FACS sorting. Such methods are described in many laboratory manuals such as, for instance, Coligan et al., Current Protocols in Immunology 1(2): Chapter 5 (1991).

For instance, expression cloning may be employed for this purpose. To this end polyadenylated RNA is prepared from a cell responsive to Lng103 and Lng104, a cDNA library is created from this RNA, the library is divided into pools and the pools are transfected individually into cells that are not responsive to Lng103 and Lng104. The transfected

cells then are exposed to labeled Lng103 and Lng104. (Lng103 and Lng104 can be labeled by a variety of well-known techniques including standard methods of radio-iodination or inclusion of a recognition site for a site-specific protein kinase.) Following exposure, the cells are fixed and binding of cytostatin is determined. These procedures conveniently are carried out on glass slides.

Pools are identified of cDNA that produced Lng103 and Lng104-binding cells. Sub-pools are prepared from these positives, transfected into host cells and screened as described above. Using an iterative sub-pooling and re-screening process, one or more single clones that encode the putative binding molecule, such as a receptor molecule, can be isolated.

Alternatively a labeled ligand can be photoaffinity linked to a cell extract, such as a membrane or a membrane extract, prepared from cells that express a molecule that it binds, such as a receptor molecule. Cross-linked material is resolved by polyacrylamide gel electrophoresis ("PAGE") and exposed to X-ray film. The labeled complex containing the ligand-receptor can be excised, resolved into peptide fragments, and subjected to protein microsequencing. The amino acid sequence obtained from microsequencing can be used to design unique or degenerate oligonucleotide probes to screen cDNA libraries to identify genes encoding the putative receptor molecule.

Polypeptides of the invention also can be used to assess Lng103 and Lng104 binding capacity of Lng103 and Lng104 binding molecules, such as receptor molecules, in cells or in cell-free preparations.

Agonists and antagonists - assays and molecules

The invention also provides a method of screening compounds to identify those which enhance or block the action of Lng103 and Lng104 on cells, such as its interaction with Lng103 and Lng104-binding molecules such

as receptor molecules. An agonist is a compound which increases the natural biological functions of Lng103 and Lng104 or which functions in a manner similar to Lng103 and Lng104, while antagonists decrease or eliminate such functions.

For example, a cellular compartment, such as a membrane or a preparation thereof, such as a membrane-preparation, may be prepared from a cell that expresses a molecule that binds Lng103 and Lng104, such as a molecule of a signaling or regulatory pathway modulated by Lng103 and Lng104. The preparation is incubated with labeled Lng103 and Lng104 in the absence or the presence of a candidate molecule which may be a Lng103 and Lng104 agonist or antagonist. The ability of the candidate molecule to bind the binding molecule is reflected in decreased binding of the labeled ligand. Molecules which bind gratuitously, i.e., without inducing the effects of HESF I, II and III on binding the Lng103 and Lng104 binding molecule, are most likely to be good antagonists. Molecules that bind well and elicit effects that are the same as or closely related to Lng103 and Lng104 are agonists. Lng103 and Lng104-like effects of potential agonists and antagonists may by measured, for instance, by determining activity of a second messenger system following interaction of the candidate molecule with a cell or appropriate cell preparation, and comparing the effect with that of Lng103 and Lng104 or molecules that elicit the same effects as Lng103 and Lng104. Second messenger systems that may be useful in this regard include, but are not limited to, AMP guanylate cyclase, ion channel or phosphoinositide hydrolysis second messenger systems.

Another example of an assay for Lng103 and Lng104 antagonists is a competitive assay that combines Lng103 and Lng104 and a potential antagonist with membrane-bound Lng103 and Lng104 receptor molecules or recombinant Lng103

and Lng104 receptor molecules under appropriate conditions for a competitive inhibition assay. Lng103 and Lng104 can be labeled, such as by radioactivity, such that the number of Lng103 and Lng104 molecules bound to a receptor molecule can be determined accurately to assess the effectiveness of the potential antagonist.

Potential antagonists include small organic molecules, peptides, polypeptides and antibodies that bind to a polypeptide of the invention and thereby inhibit or extinguish its activity. Potential antagonists also may be small organic molecules, a peptide, a polypeptide such as a closely related protein or antibody that binds the same sites on a binding molecule, such as a receptor molecule, without inducing Lng103 and Lng104-induced activities, thereby preventing the action of Lng103 and Lng104 by excluding Lng103 and Lng104 from binding.

Potential antagonists include a small molecule which binds to and occupies the binding site of the polypeptide thereby preventing binding to cellular binding molecules, such as receptor molecules, such that normal biological activity is prevented. Examples of small molecules include but are not limited to small organic molecules, peptides or peptide-like molecules.

Other potential antagonists include antisense molecules. Antisense technology can be used to control gene expression through antisense DNA or RNA or through triple-helix formation. Antisense techniques are discussed, for example, in--Okano, J. Neurochem. 56: 560 (1991); OLIGODEOXYNUCLEOTIDES AS ANTISENSE INHIBITORS OF GENE EXPRESSION, CRC Press, Boca Raton, Fla. (1988). Triple helix formation is discussed in, for instance Lee et al., Nucleic Acids Research 6: 3073 (1979); Cooney et al., Science 241: 456 (1988); and Dervan et al., Science 251: 1360 (1991). The methods are based on binding of a polynucleotide to a complementary DNA or RNA. For example,

the 5' coding portion of a polynucleotide that encodes the mature polypeptide of the present invention may be used to design an antisense RNA oligonucleotide of from about 10 to 40 base pairs in length. A DNA oligonucleotide is designed to be complementary to a region of the gene involved in transcription thereby preventing transcription and the production of Lng103 and Lng104. The antisense RNA oligonucleotide hybridizes to the mRNA *in vivo* and blocks translation of the mRNA molecule into Lng103 and Lng104 polypeptide. The oligonucleotides described above can also be delivered to cells such that the antisense RNA or DNA may be expressed *in vivo* to inhibit production of Lng103 and Lng104.

The antagonists may be employed in a composition with a pharmaceutically acceptable carrier, e.g., as hereinafter described. The antagonists may be employed for instance to treat an inherited susceptibility to asthma.

Compositions

The invention also relates to compositions comprising the polynucleotide or the polypeptides discussed above or the agonists or antagonists. Thus, the polypeptides of the present invention may be employed in combination with a non-sterile or sterile carrier or carriers for use with cells, tissues or organisms, such as a pharmaceutical carrier suitable for administration to a subject. Such compositions comprise, for instance, a media additive or a therapeutically effective amount of a polypeptide of the invention and a pharmaceutically acceptable carrier or excipient. Such carriers may include, but are not limited to, saline, buffered saline, dextrose, water, glycerol, ethanol and combinations thereof. The formulation should suit the mode of administration.

Kits

The invention further relates to pharmaceutical packs and kits comprising one or more containers filled with one

or more of the ingredients of the aforementioned compositions of the invention. Associated with such container(s) can be a notice in the form prescribed by a governmental agency regulating the manufacture, use or sale of pharmaceuticals or biological products, reflecting approval by the agency of the manufacture, use or sale of the product for human administration.

Administration

Polypeptides and other compounds of the present invention may be employed alone or in conjunction with other compounds, such as therapeutic compounds.

The pharmaceutical compositions may be administered in any effective, convenient manner including, for instance, administration by topical, oral, anal, vaginal, intravenous, intraperitoneal, intramuscular, subcutaneous, intranasal or intradermal routes among others.

The pharmaceutical compositions generally are administered in an amount effective for treatment or prophylaxis of a specific indication or indications. In general, the compositions are administered in an amount of at least about 10 μg/kg body weight. In most cases they will be administered in an amount not in excess of about 8 μg/kg body weight per day. Preferably, in most cases, dose is from about 10 μg/kg to about 2 μg/kg body weight, daily. It will be appreciated that optimum dosage will be determined by standard methods for each treatment modality and indication, taking into account the indication, its severity, route of administration, complicating conditions and the like.

Gene therapy

The Lng103 and Lng104 polynucleotides, polypeptides, agonists and antagonists that are polypeptides may be employed in accordance with the present invention by expression of such polypeptides *in vivo*, in treatment modalities often referred to as "gene therapy."

Thus, for example, cells from a patient may be engineered with a polynucleotide, such as a DNA or RNA, encoding a polypeptide *ex vivo*, and the engineered cells then can be provided to a patient to be treated with the polypeptide. For example, cells may be engineered *ex vivo* by the use of a retroviral plasmid vector containing RNA encoding a polypeptide of the present invention. Such methods are well-known in the art and their use in the present invention will be apparent from the teachings herein.

Similarly, cells may be engineered *in vivo* for expression of a polypeptide *in vivo* by procedures known in the art. For example, a polynucleotide of the invention may be engineered for expression in a replication defective retroviral vector, as discussed above. The retroviral expression construct then may be isolated and introduced into a packaging cell transduced with a retroviral plasmid vector containing RNA encoding a polypeptide of the present invention such that the packaging cell now produces infectious viral particles containing the gene of interest. These producer cells may be administered to a patient for engineering cells *in vivo* and expression of the polypeptide *in vivo*. These and other methods for administering a polypeptide of the present invention by such method should be apparent to those skilled in the art from the teachings of the present invention.

Retroviruses from which the retroviral plasmid vectors herein above mentioned may be derived include, but are not limited to, Moloney Murine Leukemia Virus, spleen necrosis virus, retroviruses such as Rous Sarcoma Virus, Harvey Sarcoma Virus, avian leukosis virus, gibbon ape leukemia virus, human immunodeficiency virus, adenovirus, Myeloproliferative Sarcoma Virus, and mammary tumor virus. In one embodiment, the retroviral plasmid vector is derived from Moloney Murine Leukemia Virus.

Such vectors will include one or more promoters for expressing the polypeptide. Suitable promoters which may be employed include, but are not limited to, the retroviral LTR; the SV40 promoter; and the human cytomegalovirus (CMV) promoter described in Miller et al., Biotechniques 7: 980-990 (1989), or any other promoter (e.g., cellular promoters such as eukaryotic cellular promoters including, but not limited to, the histone, RNA polymerase III, and .beta.-actin promoters). Other viral promoters which may be employed include, but are not limited to, adenovirus promoters, thymidine kinase (TK) promoters, and B19 parvovirus promoters. The selection of a suitable promoter will be apparent to those skilled in the art from the teachings contained herein.

The nucleic acid sequence encoding the polypeptide of the present invention will be placed under the control of a suitable promoter. Suitable promoters which may be employed include, but are not limited to, adenoviral promoters, such as the adenoviral major late promoter; or heterologous promoters, such as the cytomegalovirus (CMV) promoter; the respiratory syncytial virus (RSV) promoter; inducible promoters, such as the MMT promoter, the metallothionein promoter; heat shock promoters; the albumin promoter; the ApoAI promoter; human globin promoters; viral thymidine kinase promoters, such as the Herpes Simplex thymidine kinase promoter; retroviral LTRs (including the modified retroviral LTRs herein above described); the beta-actin promoter; and human growth hormone promoters. The promoter also may be the native promoter which controls the gene encoding the polypeptide.

The retroviral plasmid vector is employed to transduce packaging cell lines to form producer cell lines. Examples of packaging cells which may be transfected include, but are not limited to, the PE501, PA317, Y-2, Y-AM, PA12, T19-14X, VT-19-17-H2, YCRE, YCRIP, GP+E-86,

GP+envAml2, and DAN cell lines as described in Miller, A., Human Gene Therapy 1: 5-14 (1990). The vector may be transduced into the packaging cells through any means known in the art. Such means include, but are not limited to, electroporation, the use of liposomes, and CaPO4 precipitation. In one alternative, the retroviral plasmid vector may be encapsulated into a liposome, or coupled to a lipid, and then administered to a host.

The producer cell line will generate infectious retroviral vector particles, which include the nucleic acid sequence(s) encoding the polypeptides. Such retroviral vector particles then may be employed to transduce eukaryotic cells, either *in vitro* or *in vivo*. The transduced eukaryotic cells will express the nucleic acid sequence(s) encoding the polypeptide. Eukaryotic cells which may be transduced include, but are not limited to, embryonic stem cells, embryonic carcinoma cells, as well as hematopoietic stem cells, hepatocytes, fibroblasts, myoblasts, keratinocytes, endothelial cells, and bronchial epithelial cells.

EXAMPLES

The present invention is further described by the following examples. The examples are provided solely to illustrate the invention by reference to specific embodiments. These exemplification's, while illustrating certain specific aspects of the invention, do not portray the limitations or circumscribe the scope of the disclosed invention. Certain terms used herein are explained in the foregoing glossary.

All examples were carried out using standard techniques, which are well known and routine to those of skill in the art, except where otherwise described in detail. Routine molecular biology techniques of the following examples can be carried out as described in

standard laboratory manuals, such as Sambrook et al., MOLECULAR CLONING: A LABORATORY MANUAL, 2nd Ed.; Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989), herein referred to as "Sambrook."

All parts or amounts set out in the following examples are by weight, unless otherwise specified.

Unless otherwise stated size separation of fragments in the examples below was carried out using standard techniques of agarose and polyacrylamide gel electrophoresis ("PAGE") in Sambrook and numerous other references such as, for instance, by Goeddel et al., Nucleic Acids Res. 8: 4057 (1980).

Unless described otherwise, ligations were accomplished using standard buffers, incubation temperatures and times, approximately equimolar amounts of the DNA fragments to be ligated and approximately 10 units of T4 DNA ligase ("ligase") per 0.5 μg of DNA.

Example 1:Relative Quantitation of Gene Expression

Real-Time quantitative PCR with fluorescent Taqman probes is a quantitation detection system utilizing the 5'-3' nuclease activity of Taq DNA polymerase. The method uses an internal fluorescent oligonucleotide probe (Taqman) labeled with a 5' reporter dye and a downstream, 3' quencher dye. During PCR, the 5'-3' nuclease activity of Taq DNA polymerase releases the reporter, whose fluorescence can then be detected by the laser detector of the Model 7700 Sequence Detection System (PE Applied Biosystems, Foster City, CA, USA).

Amplification of an endogenous control is used to standardize the amount of sample RNA added to the reaction and normalize for Reverse Transcriptase (RT) efficiency. Either cyclophilin, glyceraldehyde-3-phosphate dehydrogenase (GAPDH), ATPase, or 18S ribosomal RNA (rRNA) is used as this endogenous control. To calculate relative

quantitation between all the samples studied, the target RNA levels for one sample were used as the basis for comparative results (calibrator). Quantitation relative to the "calibrator" can be obtained using the standard curve method or the comparative method (User Bulletin #2: ABI PRISM 7700 Sequence Detection System).

The tissue distribution and the level of the target gene were evaluated for every sample in normal and cancer tissues. Total RNA was extracted from normal tissues, cancer tissues, and from cancers and the corresponding matched adjacent tissues. Subsequently, first strand cDNA was prepared with reverse transcriptase and the polymerase chain reaction was done using primers and Taqman probes specific to each target gene. The results were analyzed using the ABI PRISM 7700 Sequence Detector. The absolute numbers are relative levels of expression of the target gene in a particular tissue compared to the calibrator tissue.

Expression of Clone ID 2798946; Gene ID 26723 (Lng103):

For the LSG Lng103, real-time quantitative PCR was performed using the following primers:

Forward Primer:

5' GGCGTTGTGGTCCTTCAG 3'(SEQ ID NO:7)

Reverse Primer

5' ACTTGCTGCCCTCTGACAC 3'(SEQ ID NO:8)

The absolute numbers depicted in Table 1 are relative levels of expression of the LSG referred to as Lng103 in 12 normal different tissues. All the values are compared to normal thymus (calibrator). These RNA samples are commercially available pools, originated by pooling samples of a particular tissue from different individuals.

Table 1: Relative Levels of LSG Lng103 Expression in Pooled Samples

TISSUE	NORMAL
Brain	0
Heart	0
Kidney	0
Liver	11
Lung	10587
Mammary Gland	19
Muscle	3
Prostate	56
Small Intestine	0
Testis	61
Thymus	1
Uterus	29

The relative levels of expression in Table 1 show that Lng103 mRNA expression is very high in lung (10587) compared with all the other normal tissues analyzed. The Lng103 mRNA expression is about 175 fold in lung tissue compared with testis tissue. These results indicate that Lng103 mRNA expression is highly specific for lung.

The absolute numbers in Table 1 were obtained analyzing pools of samples of a particular tissue from different individuals. They can not be compared to the absolute numbers originated from RNA obtained from tissue samples of a single individual in Table 2.

The absolute numbers depicted in Table 2 are relative levels of expression of Lng103 in 63 pairs of matching samples and 1 cancer and 1 normal/normal adjacent of ovary tissue. All the values are compared to normal thymus (calibrator). A matching pair is formed by mRNA from the cancer sample for a particular tissue and mRNA from the normal adjacent sample for that same tissue from the same individual.

Table 2: Relative Levels of LSG Lng103 Expression in Individual Samples

Sample ID	Cancer Type	Tissue	Cancer	Matching Normal Adjacent
LngAC82	Adenocarcinoma	Lung 1	49	198
Lng60XL	Adenocarcinoma	Lung 2	1032	8473
LngAC66	Adenocarcinoma	Lung 3	1	157
LngAC69	Adenocarcinoma	Lung 4	197	395
LngAC88	Adenocarcinoma	Lung 5	440	1050
LngAC11	Adenocarcinoma	Lung 6	50	628
LngAC39	Adenocarcinoma	Lung 7	132	12
LngAC32	Adenocarcinoma	Lung 8	144	188
LngAC90	Adenocarcinoma	Lung 9	194	249
LngAC94	Adenocarcinoma	Lung 10	74	2
LngBA641	Bronchio-alveolar carcinoma	Lung 11	235	65
LngSQ45	Squamous cell carcinoma	Lung 12	2531	251
LngSQ14	Squamous cell carcinoma	Lung 13	4	560
LngSQ9X	Squamous cell carcinoma	Lung 14	214	9
LngSQ80	Squamous cell carcinoma	Lung 15	204	7
LngSQ32	Squamous cell carcinoma	Lung 16	56	99
LngSQ16	Squamous cell carcinoma	Lung 17	20	99
LngSQ79	Squamous cell carcinoma	Lung 18	451	391
LngC20X	Squamous cell carcinoma	Lung 19	548	2463
Lng47XQ	Squamous cell carcinoma	Lung 20	15	357

LngSQ44	Squamous cell carcinoma	Lung 21	196	0
LngBR94	Squamous cell carcinoma	Lung 22	33	12
Lng90X	Squamous cell carcinoma	Lung 23	6	308
LngLC80	Large cell carcinoma	Lung 24	25	215
LngLC71	Large cell carcinoma	Lung 25	895	299
LngLC109	Large cell carcinoma	Lung 26	5	118
LngMT67	Metastatic from renal cell cancer	Lung 27	1	6
LngMT71	Metastatic from melanoma	Lung 28	0	206
Bld32XK		Bladder 1	0	0
Bld66X		Bladder 2	0	0
CvxKS52		Cervix 1	0	0
CvxNK23		Cervix 2	54	0
ClnAS45		Colon 1	0	0
ClnRC24		Colon 2	0	0
Endo28XA		Endometrium 1	3	0
Endo68X		Endometrium 2	1	2
Endo10479		Endometrium 3	2	0
Kid106XD		Kidney 1	0	0
Kid109XD		Kidney 2	0	0
Liv94XA		Liver 1	0	0
Liv15XA		Liver 2	0	0
Mam162X		Mammary Gland 1	9	1

Mam19DN		Mammary Gland 2	14	1
MamA06X		Mammary Gland 3	1	0
Ovr108B		Ovary 1	0	-
Ovr103X		Ovary 2	0	0
Ovr18GA		Ovary 3	-	0
Pan71XL		Pancreas 1	0	0
Pan92X		Pancreas 2	122	0
Pro90XB		Prostate 1	31	4
Pro101XB		Prostate 2	2	1
Skn248S		Skin 1	0	0
SmIntH89		Small Intestine 1	0	0
StoAC44		Stomach 1	459	15
StoAC93		Stomach 2	459	86
Sto115S		Stomach 3	0	381
Sto531S		Stomach 4	0	2
Sto261S		Stomach 5	173	35
Tst39X		Testis 1	29	0
Tst647T		Testis 2	31	0
Tst663T		Testis 3	1	0
Thr145T		Thyroid 1	1	0
Thy143N		Thymus 1	0	0
Utr135XO		Uterus 1	0	0
Utr141XO		Uterus 2	1	1

0 = Negative

In the analysis of matching samples, the higher levels of expression were in lung showing a high degree of tissue specificity for lung tissue. These results confirm the tissue specificity results obtained with normal pooled samples (Table 1).

Furthermore, the level of mRNA expression in cancer samples and the isogenic normal adjacent tissue from the same individual were compared. This comparison provides an indication of specificity for the cancer stage (e.g. higher levels of mRNA expression in the cancer sample compared to the normal adjacent). Table 2 shows overexpression of Lng103 in 10 lung cancer tissues compared with their respective normal adjacent (lung samples #7, 10, 11, 12, 14, 15, 18, 21, 22, and 25). Thus, there was overexpression in the cancer tissue for 36% of the lung matching samples tested (total of 28 lung matching samples).

The absolute numbers depicted in Table 3 show the relative levels of expression of Lng103 in 12 pairs of matching samples of lung squamous cell carcinoma.

Table 3: Relative Levels of LSG Lng103 Expression in Matching Sample Pairs

Sample ID	Cancer Type	Tissue	Cancer	Matching Normal Adjacent
LngSQ45	Squamous cell carcinoma	Lung 1	2531	251
LngSQ14	Squamous cell carcinoma	Lung 2	4	560
LngSQ9X	Squamous cell carcinoma	Lung 3	214	9
LngSQ80	Squamous cell carcinoma	Lung 4	204	7
LngSQ32	Squamous cell carcinoma	Lung 5	56	99
LngSQ16	Squamous cell carcinoma	Lung 6	20	99
LngSQ79	Squamous cell carcinoma	Lung 7	451	391
LngC20X	Squamous cell carcinoma	Lung 8	548	2463

Lng47XQ	Squamous cell carcinoma	Lung 9	15	357
LngSQ44	Squamous cell carcinoma	Lung 10	196	0
LngBR94	Squamous cell carcinoma	Lung 11	33	12
Lng90X	Squamous cell carcinoma	Lung 12	6	308

Table 3 shows overexpression of Lng103 in 6 cancer tissues compared with their respective normal adjacent (lung samples #1, 3, 4, 7, 10, and 11). Thus, there is overexpression of Lng103 in the cancer tissue for 50% of the lung squamous cell carcinoma (total 12 matching pairs).

Altogether, the high level of tissue specificity, plus the mRNA overexpression in 50% of the lung squamous cell carcinoma matching samples tested are indicative of Lng103 being a diagnostic marker for lung cancer, particularly squamous cell carcinoma.

Expression of Clone ID 126263; Gene ID 221807 (Lng104):

For the LSG Lng104, real-time quantitative PCR was performed using the following primers:

Forward Primer

5' ATCACAGGGGCACTGCTTCT-3'(SEQ ID NO:9)

Reverse Primer

5'-GACCGAGGCCAGGCTTCTA-3'(SEQ ID NO:10)

The absolute numbers depicted in Table 4 are relative levels of expression of the LSG Lng104 in 12 normal different tissues. All the values are compared to normal small intestine (calibrator). These RNA samples are commercially available pools, originated by pooling samples of a particular tissue from different individuals.

Table 4: Relative Levels of LSG Lng104 Expression in Pooled Samples

TISSUE	NORMAL
Brain	0
Heart	0
Kidney	2
Liver	0
Lung	7617
Mammary Gland	113
Muscle	5
Prostate	171
Small Intestine	1
Testis	3
Thymus	8
Uterus	11

The relative levels of expression in Table 4 show that Lng104 mRNA expression is very high (7617) in lung compared with all the other normal tissues analyzed. The Lng104 mRNA expression is about 45 fold in lung tissue compared with prostate tissue. These results demonstrated that Lng104 mRNA expression is highly specific for lung.

The absolute numbers in Table 4 were obtained analyzing pools of samples of a particular tissue from different individuals. They can not be compared to the absolute numbers originated from RNA obtained from tissue samples of a single individual in Table 5.

The absolute numbers depicted in Table 5 are relative levels of expression of Lng104 in 64 pairs of matching samples and 1 cancer and 1 normal/normal adjacent of ovary tissue. All the values are compared to normal testis (calibrator). A matching pair is formed by mRNA from the cancer sample for a particular tissue and mRNA from the normal adjacent sample for that same tissue from the same individual.

Table 5: Relative Levels of LSG Lng104 Expression in Individual Samples

Sample ID	Cancer Type	Tissue	Cancer	Matching Normal Adjacent
LngAC82	Adenocarcinoma	Lung 1	4111	28333
Lng60XL	Adenocarcinoma	Lung 2	244	7887
LngAC66	Adenocarcinoma	Lung 3	2927	23991
LngAC69	Adenocarcinoma	Lung 4	13081	24241
LngAC88	Adenocarcinoma	Lung 5	2304	27270
LngAC11	Adenocarcinoma	Lung 6	503	12232
LngAC39	Adenocarcinoma	Lung 7	8800	8842
LngAC32	Adenocarcinoma	Lung 8	3020	11346
LngAC90	Adenocarcinoma	Lung 9	5257	14412
LngAC94	Adenocarcinoma	Lung 10	7156	8422
LngBA641	Bronchio-alveolar carcinoma	Lung 11	25709	22304
LngSQ9X	Squamous cell carcinoma	Lung 12	25009	18757
LngSQ45	Squamous cell carcinoma	Lung 13	21552	1841
LngSQ14	Squamous cell carcinoma	Lung 14	509	13523
LngSQ32	Squamous cell carcinoma	Lung 15	1292	19336
LngSQ80	Squamous cell carcinoma	Lung 16	204	7
LngSQ16	Squamous cell carcinoma	Lung 17	5542	3754
LngSQ79	Squamous cell carcinoma	Lung 18	1365	10662
Lng47XQ	Squamous cell carcinoma	Lung 19	11388	23660
LngC20X	Squamous cell carcinoma	Lung 20	16997	5583

LngSQ44	Squamous cell carcinoma	Lung 21	4798	233
LngBR94	Squamous cell carcinoma	Lung 22	33	12
Lng90X	Squamous cell carcinoma	Lung 23	570	3386
LngLC80	Large cell carcinoma	Lung 24	5631	10886
LngLC71	Large cell carcinoma	Lung 25	19690	19827
LngLC109	Large cell carcinoma	Lung 26	14666	41338
LngMT67	Metastatic from renal cell cancer	Lung 27	117	13449
LngMT71	Metastatic from melanoma	Lung 28	615	6818
Bld32XK		Bladder 1	349	1214
Bld66X		Bladder 2	2740	1337
ClnAS45		Colon 1	1	0
ClnRC24		Colon 2	5	0
CvxKS52		Cervix 1	7282	3154
CvxNK23		Cervix 2	15614	7602
Endo10479		Endometrium 1	131	24
Endo12XA		Endometrium 2	823	172
Endo8911		Endometrium 3	1967	76
Endo28XA		Endometrium 4	22152	67
Endo68X		Endometrium 5	183	190
Kid106XD		Kidney 1	0	24
Kid109XD		Kidney 2	15	2
Liv15XA		Liver 1	2	1

Liv94XA		Liver 2	0	0
Mam162X		Mammary Gland 1	51	13
Mam19DN		Mammary Gland 2	5	177
MamA06X		Mammary Gland 3	151	69
Ovr103X		Ovary 1	299	0
Ovr180B		Ovary 2	30	-
Ovr18GA		Ovary 3	-	2
Pan71XL		Pancreas 1	17	9
Pro20XB		Prostate 1	73	1607
Pro90XB		Prostate 2	58	28
Pro101XB		Prostate 3	83	157
Skn248S		Skin 1	6	1
SmInt21XA		Small Intestine 1	3	0
SmIntH89		Small Intestine 2	4	0
Sto115S		Stomach 1	60	17805
Sto261S		Stomach 2	589	1592
Sto531S		Stomach 3	2	5803
StoAC44		Stomach 4	273	8813
StoAC93		Stomach 5	4683	7767
Tst39X		Testis 1	541	4
Tst647T		Testis 2	203	0
Thy143N		Thymus 1	169	16
Utr135XO		Uterus 1	74	73
Utr141XO		Uterus 2	82	3

0 = Negative

In the analysis of matching samples, the higher levels of expression were in lung showing a high degree of tissue specificity for lung tissue. These results confirm

the tissue specificity results obtained with normal pooled samples (Table 4).

Furthermore, the level of mRNA expression in cancer samples and the isogenic normal adjacent tissue from the same individual were compared. This comparison provides an indication of specificity for the cancer stage (e.g. higher levels of mRNA expression in the cancer sample compared to the normal adjacent). Table 5 shows overexpression of Lng104 in 8 lung cancer tissues compared with their respective normal adjacent (lung samples #11, 12, 13, 16, 17, 20, 21, and 22). Thus, there is overexpression in the cancer tissue for 28% of the lung matching samples tested (total of 28 lung matching samples).

The absolute numbers depicted in Table 6 are relative levels of expression of Lng104 in 12 pairs of matching samples of lung squamous cell carcinoma.

Table 6: Relative Levels of LSG Lng104 Expression in Matching Sample Pairs

Sample ID	Cancer Type	Tissue	Cancer	Matching Normal Adjacent
LngSQ45	Squamous cell carcinoma	Lung 1	25009	18757
LngSQ14	Squamous cell carcinoma	Lung 2	21552	1841
LngSQ9X	Squamous cell carcinoma	Lung 3	509	13523
LngSQ80	Squamous cell carcinoma	Lung 4	1292	19336
LngSQ32	Squamous cell carcinoma	Lung 5	204	7
LngSQ16	Squamous cell carcinoma	Lung 6	5542	3754
LngSQ79	Squamous cell carcinoma	Lung 7	1365	10662

LngC20X	Squamous cell carcinoma	Lung 8	11388	23660
Lng47XQ	Squamous cell carcinoma	Lung 9	16997	5583
LngSQ44	Squamous cell carcinoma	Lung 10	4798	233
LngBR94	Squamous cell carcinoma	Lung 11	33	12
Lng90X	Squamous cell carcinoma	Lung 12	570	3386

Table 6 shows overexpression of Lng104 in 7 cancer tissues compared with their respective normal adjacent (lung samples #1, 2, 5, 6, 9, 10 and 11). There is overexpression of Lng104 in the cancer tissue for 58% of the lung squamous cell carcinoma (total 12 matching pairs).

Altogether, the high level of tissue specificity, plus the mRNA overexpression in 58% of the lung squamous cell carcinoma matching samples tested are indicative of Lng104 being a diagnostic marker for lung cancer, particularly squamous cell carcinoma.

Example 2: Protein Expression

Lng104 was amplified by polymerase chain reaction (PCR) and the amplified DNA fragment encoding amino acid number 16 (Leu^{16}) to amino acid number 119 (Leu^{119}) of Lng104 was subcloned in pET-21d for expression in *E. coli*. In addition to Lng104 DNA coding sequence (Leu^{16}- Leu^{119}), codons for two amino acids, Met-Ala, flanking the NH2-terminus of Lng104, and six histidines, flanking the COOH-terminus of Lng104, were incorporated to serve as initiating Met/restriction site and purification tag, respectively. An over-expressed protein band of approximately 13 kDa (translated molecular weight, 13.2 kDa) was readily observed on a Coomassie blue stained polyacrylamide gel. This protein band was confirmed by Western blot analysis using monoclonal antibody against 6 x Histidine tag.

Large-scale purification of Lng104 was achieved using cell paste generated from 6-liter bacterial cultures, and purified using immobilized metal affinity chromatography (IMAC). Soluble fraction separated from total cell lysate was incubated with a nickle chelating resin. Column was packed and washed with five column volumes of wash buffer. Lng104 was eluted stepwise with various concentration imidazole buffers.

In a similar manner, Lng103 is expressed and purified.

Example 3: Protein Fusions

The polypeptides of the present invention are preferably fused to other proteins. These fusion proteins can be used for a variety of applications. For example, fusion of the present polypeptides to His-tag, HA-tag, protein A, IgG domains, and maltose binding protein facilitates purification. (See also EP A 394,827; Traunecker, et al., Nature 331: 84-86 (1988).) Similarly, fusion to IgG-1, IgG-3, and albumin increases the halflife time *in vivo*. Nuclear localization signals fused to the polypeptides of the present invention can target the protein to a specific subcellular localization, while covalent heterodimer or homodimers can increase or decrease the activity of a fusion protein. Fusion proteins can also create chimeric molecules having more than one function. Finally, fusion proteins can increase solubility and/or stability of the fused protein compared to the non-fused protein. All of the types of fusion proteins described above can be made by modifying the following protocol, which outlines the fusion of a polypeptide to an IgG molecule, or the protocol described above.

Briefly, the human Fc portion of the IgG molecule (SEQ ID NO:11) can be PCR amplified, using primers that span the 5'and 3'ends of the sequence described below. These primers also should have convenient restriction enzyme sites that will facilitate cloning into an expression vector, preferably a mammalian expression

vector.

For example, if pC4 (Accession No. 209646) is used, the human Fc portion can be ligated into the BamHI cloning site. Note that the 3'BamHI site should be destroyed. Next, the vector containing the human Fc portion is re-restricted with BamHI, linearizing the vector, and a polynucleotide of the present invention, isolated by the PCR protocol described in Example 1, is ligated into this BamHI site. Note that the polynucleotide is cloned without a stop codon, otherwise a fusion protein will not be produced.

If the naturally occurring signal sequence is used to produce the secreted protein, pC4 does not need a second signal peptide. Alternatively, if the naturally occurring signal sequence is not used, the vector can be modified to include a heterologous signal sequence. (See, e. g., WO 96/34891.)

Human IgG Fc region is depicted in SEQ ID NO:11.

Example 4: Production of an Antibody from a Polypeptide

The antibodies of the present invention can be prepared by a variety of methods. (See, Current Protocols, Chapter 2.) For example, cells expressing a polypeptide of the present invention is administered to an animal to induce the production of sera containing polyclonal antibodies. In a preferred method, a preparation of the secreted protein is prepared and purified to render it substantially free of natural contaminants. Such a preparation is then introduced into an animal in order to produce polyclonal antisera of greater specific activity.

In the most preferred method, the antibodies of the present invention are monoclonal antibodies (or protein binding fragments thereof). Such monoclonal antibodies can be prepared using hybridoma technology. (Kohler et al., Nature 256: 495 (1975); Khler et al., Eur. J. Immunol. 6: 511 (1976); Khler et al., Eur. J.Immunol. 6: 292 (1976); Hammerling et al., in: Monoclonal Antibodies and T-Cell

Hybridomas, Elsevier, N. Y., pp. 563-681 (1981).) In general, such procedures involve immunizing an animal (preferably a mouse) with polypeptide or, more preferably, with a secreted polypeptide-expressing cell. Such cells may be cultured in any suitable tissue culture medium; however, it is preferable to culture cells in Earle's modified Eagle's medium supplemented with 10% fetal bovine serum (inactivated at about 56°C), and supplemented with about 10 g/l of nonessential amino acids, about 1,000 U/ml of penicillin, and about 100, μg/ml of streptomycin.

The splenocytes of such mice are extracted and fused with a suitable myeloma cell line. Any suitable myeloma cell line may be employed in accordance with the present invention; however, it is preferable to employ the parent myeloma cell line (SP20), available from the ATCC. After fusion, the resulting hybridoma cells are selectively maintained in HAT medium, and then cloned by limiting dilution as described by Wands et al. (Gastroenterology 80: 225-232 (1981).) The hybridoma cells obtained through such a selection are then assayed to identify clones which secrete antibodies capable of binding the polypeptide.

Alternatively, additional antibodies capable of binding to the polypeptide can be produced in a two-step procedure using anti-idiotypic antibodies. Such a method makes use of the fact that antibodies are themselves antigens, and therefore, it is possible to obtain an antibody which binds to a second antibody. In accordance with this method, protein specific antibodies are used to immunize an animal, preferably a mouse. The splenocytes of such an animal are then used to produce hybridoma cells, and the hybridoma cells are screened to identify clones which produce an antibody whose ability to bind to the protein-specific antibody can be blocked by the polypeptide. Such antibodies comprise anti-idiotypic antibodies to the protein specific antibody and can be used

to immunize an animal to induce formation of further protein-specific antibodies.

It will be appreciated that Fab and F (ab') 2 and other fragments of the antibodies of the present invention may be used according to the methods disclosed herein. Such fragments are typically produced by proteolytic cleavage, using enzymes such as papain (to produce Fab fragments) or pepsin (to produce F (ab') 2 fragments). Alternatively, secreted protein-binding fragments can be produced through the application of recombinant DNA technology or through synthetic chemistry.

For *in vivo* use of antibodies in humans, it may be preferable to use "humanized"chimeric monoclonal antibodies. Such antibodies can be produced using genetic constructs derived from hybridoma cells producing the monoclonal antibodies described above. Methods for producing chimeric antibodies are known in the art. (See, for review, Morrison, Science 229: 1202 (1985); Oi et al., BioTechniques 4: 214 (1986); Cabilly et al., U. S. Patent No. 4,816,567; Taniguchi et al., EP 171496; Morrison et al., EP 173494; Neuberger et al., WO 8601533; Robinson et al., WO 8702671; Boulianne et al., Nature 312: 643 (1984); Neuberger et al., Nature 314: 268 (1985).)

Example 5: Method of Determining Alterations in a Gene Corresponding to a Polynucleotide

RNA isolated from entire families or individual patients presenting with a phenotype of interest (such as a disease) is be isolated. cDNA is then generated from these RNA samples using protocols known in the art. (See, Sambrook.) The cDNA is then used as a template for PCR, employing primers surrounding regions of interest in SEQ ID NO: 1, 2, 5 and 6. Suggested PCR conditions consist of 35 cycles at 95°C for 30 seconds; 60-120 seconds at 52-58°C; and 60-120 seconds at 70°C, using buffer solutions described in Sidransky, D., et al., Science 252: 706 (1991).

PCR products are then sequenced using primers labeled at their 5'end with T4 polynucleotide kinase, employing SequiTherm Polymerase. (Epicentre Technologies). The intron-exon borders of selected exons is also determined and genomic PCR products analyzed to confirm the results. PCR products harboring suspected mutations is then cloned and sequenced to validate the results of the direct sequencing.

PCR products is cloned into T-tailed vectors as described in Holton, T. A. and Graham, M. W., Nucleic Acids Research, 19 : 1156 (1991) and sequenced with T7 polymerase (United States Biochemical). Affected individuals are identified by mutations not present in unaffected individuals.

Genomic rearrangements are also observed as a method of determining alterations in a gene corresponding to a polynucleotide. Genomic clones are nick-translated with digoxigenin deoxy-uridine 5'triphosphate (Boehringer Manheim), and FISH performed as described in Johnson, C. et al., Methods Cell Biol. 35: 73-99 (1991). Hybridization with the labeled probe is carried out using a vast excess of human cot-1 DNA for specific hybridization to the corresponding genomic locus.

Chromosomes are counterstained with 4,6-diamino-2-phenylidole and propidium iodide, producing a combination of C-and R-bands. Aligned images for precise mapping are obtained using a triple-band filter set (Chroma Technology, Brattleboro, VT) in combination with a cooled charge-coupled device camera (Photometrics, Tucson, AZ) and variable excitation wavelength filters. (Johnson, Cv. et al., Genet. Anal. Tech. Appl., 8: 75 (1991).) Image collection, analysis and chromosomal fractional length measurements are performed using the ISee Graphical Program System. (Inovision Corporation, Durham, NC.) Chromosome alterations of the genomic region hybridized by

the probe are identified as insertions, deletions, and translocations. These alterations are used as a diagnostic marker for an associated disease.

Example 6: Method of Detecting Abnormal Levels of a Polypeptide in a Biological Sample

A polypeptide of the present invention can be detected in a biological sample, and if an increased or decreased level of the polypeptide is detected, this polypeptide is a marker for a particular phenotype. Methods of detection are numerous, and thus, it is understood that one skilled in the art can modify the following assay to fit their particular needs.

For example, antibody-sandwich ELISAs are used to detect polypeptides in a sample, preferably a biological sample. Wells of a microtiter plate are coated with specific antibodies, at a final concentration of 0.2 to 10 ug/ml. The antibodies are either monoclonal or polyclonal and are produced by the method described above. The wells are blocked so that non-specific binding of the polypeptide to the well is reduced.

The coated wells are then incubated for > 2 hours at RT with a sample containing the polypeptide. Preferably, serial dilutions of the sample should be used to validate results. The plates are then washed three times with deionized or distilled water to remove unbounded polypeptide.

Next, 50 ul of specific antibody-alkaline phosphatase conjugate, at a concentration of 25-400 ng, is added and incubated for 2 hours at room temperature.

The plates are again washed three times with deionized or distilled water to remove unbounded conjugate.

Add 75 ul of 4-methylumbelliferyl phosphate (MUP) or p-nitrophenyl phosphate (NPP) substrate solution to each well and incubate 1 hour at room temperature. Measure the reaction by a microtiter plate reader. Prepare a standard

curve, using serial dilutions of a control sample, and plot polypeptide concentration on the X-axis (log scale) and fluorescence or absorbance of the Y-axis (linear scale).

Interpolate the concentration of the polypeptide in the sample using the standard curve.

Example 7: Formulating a Polypeptide

The secreted polypeptide composition will be formulated and dosed in a fashion consistent with good medical practice, taking into account the clinical condition of the individual patient (especially the side effects of treatment with the secreted polypeptide alone), the site of delivery, the method of administration, the scheduling of administration, and other factors known to practitioners. The "effective amount" for purposes herein is thus determined by such considerations.

As a general proposition, the total pharmaceutically effective amount of secreted polypeptide administered parenterally per dose will be in the range of about 1 , μg/kg/day to 10 mg/kg/day of patient body weight, although, as noted above, this will be subject to therapeutic discretion. More preferably, this dose is at least 0.01 mg/kg/day, and most preferably for humans between about 0.01 and 1 mg/kg/day for the hormone. If given continuously, the secreted polypeptide is typically administered at a dose rate of about 1 μg/kg/hour to about 50 mg/kg/hour, either by 1-4 injections per day or by continuous subcutaneous infusions, for example, using a mini-pump. An intravenous bag solution may also be employed. The length of treatment needed to observe changes and the interval following treatment for responses to occur appears to vary depending on the desired effect.

Pharmaceutical compositions containing the secreted protein of the invention are administered orally, rectally, parenterally, intracistemally, intravaginally, intraperitoneally, topically (as by powders, ointments,

gels, drops or transdermal patch), bucally, or as an oral or nasal spray. "Pharmaceutically acceptable carrier" refers to a non-toxic solid, semisolid or liquid filler, diluent, encapsulating material or formulation auxiliary of any type. The term "parenteral" as used herein refers to modes of administration which include intravenous, intramuscular, intraperitoneal, intrasternal, subcutaneous and intraarticular injection and infusion.

The secreted polypeptide is also suitably administered by sustained-release systems. Suitable examples of sustained-release compositions include semipermeable polymer matrices in the form of shaped articles, e. g., films, or microcapsules. Sustained-release matrices include polylactides (U. S. Pat. No.3,773,919, EP 58,481), copolymers of L-glutamic acid and gamma-ethyl-L-glutamate (Sidman, U. et al., Biopolymers 22: 547-556 (1983)), poly (2-hydroxyethyl methacrylate) (R. Langer et al., J. Biomed. Mater. Res. 15: 167-277 (1981), and R. Langer, Chem. Tech. 12: 98-105 (1982)), ethylene vinyl acetate (R. Langer et al.) or poly-D- (-)-3-hydroxybutyric acid (EP 133,988). Sustained-release compositions also include liposomally entrapped polypeptides. Liposomes containing the secreted polypeptide are prepared by methods known per se: DE Epstein et al., Proc. Natl. Acad. Sci. USA 82: 3688-3692 (1985); Hwang et al., Proc. Natl. Acad. Sci. USA 77: 4030-4034 (1980); EP 52,322; EP 36,676; EP 88,046; EP 143,949; EP 142,641; Japanese Pat. Appl. 83-118008; U. S. Pat. Nos. 4,485,045 and 4,544,545; and EP 102,324. Ordinarily, the liposomes are of the small (about 200-800 Angstroms) unilamellar type in which the lipid content is greater than about 30 mol. percent cholesterol, the selected proportion being adjusted for the optimal secreted polypeptide therapy.

For parenteral administration, in one embodiment, the secreted polypeptide is formulated generally by mixing it

at the desired degree of purity, in a unit dosage injectable form (solution, suspension, or emulsion), with a pharmaceutically acceptable carrier, I. e., one that is non-toxic to recipients at the dosages and concentrations employed and is compatible with other ingredients of the formulation.

For example, the formulation preferably does not include oxidizing agents and other compounds that are known to be deleterious to polypeptides.

Generally, the formulations are prepared by contacting the polypeptide uniformly and intimately with liquid carriers or finely divided solid carriers or both.

Then, if necessary, the product is shaped into the desired formulation. Preferably the carrier is a parenteral carrier, more preferably a solution that is isotonic with the blood of the recipient. Examples of such carrier vehicles include water, saline, Ringer's solution, and dextrose solution. Non-aqueous vehicles such as fixed oils and ethyl oleate are also useful herein, as well as liposomes.

The carrier suitably contains minor amounts of additives such as substances that enhance isotonicity and chemical stability. Such materials are non-toxic to recipients at the dosages and concentrations employed, and include buffers such as phosphate, citrate, succinate, acetic acid, and other organic acids or their salts; antioxidants such as ascorbic acid; low molecular weight (less than about ten residues) polypeptides, e. g., polyarginine or tripeptides; proteins, such as serum albumin, gelatin, or immunoglobulins; hydrophilic polymers such as polyvinylpyrrolidone; amino acids, such as glycine, glutamic acid, aspartic acid, or arginine; monosaccharides, disaccharides, and other carbohydrates including cellulose or its derivatives, glucose, manose, or dextrins; chelating agents such as EDTA; sugar alcohols such as mannitol or

sorbitol; counterions such as sodium; and/or nonionic surfactants such as polysorbates, poloxamers, or PEG.

The secreted polypeptide is typically formulated in such vehicles at a concentration of about 0.1 mg/ml to 100 mg/ml, preferably 1-10 mg/ml, at a pH of about 3 to 8. It will be understood that the use of certain of the foregoing excipients, carriers, or stabilizers will result in the formation of polypeptide salts.

Any polypeptide to be used for therapeutic administration can be sterile. Sterility is readily accomplished by filtration through sterile filtration membranes (e. g., 0.2 micron membranes). Therapeutic polypeptide compositions generally are placed into a container having a sterile access port, for example, an intravenous solution bag or vial having a stopper pierceable by a hypodermic injection needle.

Polypeptides ordinarily will be stored in unit or multi-dose containers, for example, sealed ampules or vials, as an aqueous solution or as a lyophilized formulation for reconstitution. As an example of a lyophilized formulation, 10-ml vials are filled with 5 ml of sterile-filtered 1 % (w/v) aqueous polypeptide solution, and the resulting mixture is lyophilized. The infusion solution is prepared by reconstituting the lyophilized polypeptide using bacteriostatic Water-for-Injection.

The invention also provides a pharmaceutical pack or kit comprising one or more containers filled with one or more of the ingredients of the pharmaceutical compositions of the invention. Associated with such container (s) can be a notice in the form prescribed by a governmental agency regulating the manufacture, use or sale of pharmaceuticals or biological products, which notice reflects approval by the agency of manufacture, use or sale for human administration. In addition, the polypeptides of the present invention may be employed in conjunction with other

therapeutic compounds.

Example 8: Method of Treating Decreased Levels of the Polypeptide

It will be appreciated that conditions caused by a decrease in the standard or normal expression level of a secreted protein in an individual can be treated by administering the polypeptide of the present invention, preferably in the secreted form. Thus, the invention also provides a method of treatment of an individual in need of an increased level of the polypeptide comprising administering to such an individual a pharmaceutical composition comprising an amount of the polypeptide to increase the activity level of the polypeptide in such an individual.

For example, a patient with decreased levels of a polypeptide receives a daily dose 0.1-100 ug/kg of the polypeptide for six consecutive days. Preferably, the polypeptide is in the secreted form. The exact details of the dosing scheme, based on administration and formulation, are provided above.

Example 9: Method of Treating Increased Levels of the Polypeptide

Antisense technology is used to inhibit production of a polypeptide of the present invention. This technology is one example of a method of decreasing levels of a polypeptide, preferably a secreted form, due to a variety of etiologies, such as cancer.

For example, a patient diagnosed with abnormally increased levels of a polypeptide is administered intravenously antisense polynucleotides at 0.5, 1.0, 1.5, 2.0 and 3.0 mg/kg day for 21 days. This treatment is repeated after a 7-day rest period if the treatment was well tolerated. The formulation of the antisense polynucleotide is provided above.

Example 10: Method of Treatment Using Gene Therapy

One method of gene therapy transplants fibroblasts, which are capable of expressing a polypeptide, onto a patient. Generally, fibroblasts are obtained from a subject by skin biopsy. The resulting tissue is placed in tissue-culture medium and separated into small pieces. Small chunks of the tissue are placed on a wet surface of a tissue culture flask, approximately ten pieces are placed in each flask. The flask is turned upside down, closed tight and left at room temperature over night. After 24 hours at room temperature, the flask is inverted and the chunks of tissue remain fixed to the bottom of the flask and fresh media (e. g., Ham's F12 media, with 10% FBS, penicillin and streptomycin) is added. The flasks are then incubated at 37°C for approximately one week.

At this time, fresh media is added and subsequently changed every several days. After an additional two weeks in culture, a monolayer of fibroblasts emerge. The monolayer is trypsinized and scaled into larger flasks. pMV-7 (Kirschmeier, P. T. et al., DNA, 7: 219-25 (1988)), flanked by the long terminal repeats of the Moloney murine sarcoma virus, is digested with EcoRI and HindIII and subsequently treated with calf intestinal phosphatase. The linear vector is fractionated on agarose gel and purified, using glass beads.

The cDNA encoding a polypeptide of the present invention can be amplified using PCR primers which correspond to the 5'and 3'end sequences respectively as set forth in Example 1. Preferably, the 5'primer contains an EcoRI site and the 3'primer includes a HindIII site. Equal quantities of the Moloney murine sarcoma virus linear backbone and the amplified EcoRI and HindIII fragment are added together, in the presence of T4 DNA ligase. The resulting mixture is maintained under conditions appropriate for ligation of the two fragments. The ligation

mixture is then used to transform bacteria HB 101, which are then plated onto agar containing kanamycin for the purpose of confirming that the vector has the gene of interest properly inserted.

The amphotropic pA317 or GP+aml2 packaging cells are grown in tissue culture to confluent density in Dulbecco's Modified Eagles Medium (DMEM) with 10% calf serum (CS), penicillin and streptomycin. The MSV vector containing the gene is then added to the media and the packaging cells transduced with the vector. The packaging cells now produce infectious viral particles containing the gene (the packaging cells are now referred to as producer cells).

Fresh media is added to the transduced producer cells, and subsequently, the media is harvested from a 10 cm plate of confluent producer cells. The spent media, containing the infectious viral particles, is filtered through a millipore filter to remove detached producer cells and this media is then used to infect fibroblast cells. Media is removed from a sub-confluent plate of fibroblasts and quickly replaced with the media from the producer cells. This media is removed and replaced with fresh media.

If the titer of virus is high, then virtually all fibroblasts will be infected and no selection is required. If the titer is very low, then it is necessary to use a retroviral vector that has a selectable marker, such as neo or his. Once the fibroblasts have been efficiently infected, the fibroblasts are analyzed to determine whether protein is produced.

The engineered fibroblasts are then transplanted onto the host, either alone or after having been grown to confluence on cytodex 3 microcarrier beads.

Example 11: Method of Treatment Using Gene Therapy-*In Vivo*

Another aspect of the present invention is using *in vivo* gene therapy methods to treat disorders, diseases and

conditions. The gene therapy method relates to the introduction of naked nucleic acid (DNA, RNA, and antisense DNA or RNA) sequences into an animal to increase or decrease the expression of the polypeptide.

The polynucleotide of the present invention may be operatively linked to a promoter or any other genetic elements necessary for the expression of the polypeptide by the target tissue. Such gene therapy and delivery techniques and methods are known in the art, see, for example, WO 90/11092, WO 98/11779; U. S. Patent No. 5,693,622; 5,705,151; 5,580,859; Tabata H. et al. (1997) Cardiovasc. Res. 35 (3): 470-479, Chao J et al. (1997) Pharmacol. Res. 35 (6): 517-522, Wolff J. A. (1997) Neuromuscul. Disord. 7 (5): 314-318, Schwartz B. et al. (1996) Gene Ther. 3 (5): 405-411, Tsurumi Y. et al. (1996) Circulation 94 (12): 3281-3290 (incorporated herein by reference).

The polynucleotide constructs may be delivered by any method that delivers injectable materials to the cells of an animal, such as, injection into the interstitial space of tissues (heart, muscle, skin, lung, liver, intestine and the like). The polynucleotide constructs can be delivered in a pharmaceutically acceptable liquid or aqueous carrier.

The term "naked" polynucleotide, DNA or RNA, refers to sequences that are free from any delivery vehicle that acts to assist, promote, or facilitate entry into the cell, including viral sequences, viral particles, liposome formulations, lipofectin or precipitating agents and the like. However, the polynucleotides of the present invention may also be delivered in liposome formulations (such as those taught in Felgner P. L. et al. (1995) Ann. NY Acad. Sci. 772: 126-139 and Abdallah B. et al. (1995) Biol. Cell 85 (1): 1-7) which can be prepared by methods well known to those skilled in the art.

The polynucleotide vector constructs used in the gene

therapy method are preferably constructs that will not integrate into the host genome nor will they contain sequences that allow for replication. Any strong promoter known to those skilled in the art can be used for driving the expression of DNA. Unlike other gene therapies techniques, one major advantage of introducing naked nucleic acid sequences into target cells is the transitory nature of the polynucleotide synthesis in the cells. Studies have shown that non-replicating DNA sequences can be introduced into cells to provide production of the desired polypeptide for periods of up to six months.

The polynucleotide construct can be delivered to the interstitial space of tissues within the an animal, including of muscle, skin, brain, lung, liver, spleen, bone marrow, thymus, heart, lymph, blood, bone, cartilage, pancreas, kidney, gall bladder, stomach, intestine, testis, ovary, uterus, rectum, nervous system, eye, gland, and connective tissue. Interstitial space of the tissues comprises the intercellular fluid, mucopolysaccharide matrix among the reticular fibers of organ tissues, elastic fibers in the walls of vessels or chambers, collagen fibers of fibrous tissues, or that same matrix within connective tissue ensheathing muscle cells or in the lacunae of bone. It is similarly the space occupied by the plasma of the circulation and the lymph fluid of the lymphatic channels. Delivery to the interstitial space of muscle tissue is preferred for the reasons discussed below. They may be conveniently delivered by injection into the tissues comprising these cells. They are preferably delivered to and expressed in persistent, non-dividing cells which are differentiated, although delivery and expression may be achieved in non-differentiated or less completely differentiated cells, such as, for example, stem cells of blood or skin fibroblasts. *In vivo* muscle cells are particularly competent in their ability to take up and

express polynucleotides.

For the naked polynucleotide injection, an effective dosage amount of DNA or RNA will be in the range of from about 0.05 μg/kg body weight to about 50 mg/kg body weight. Preferably the dosage will be from about 0.005 mg/kg to about 20 mg/kg and more preferably from about 0.05 mg/kg to about 5 mg/kg. Of course, as the artisan of ordinary skill will appreciate, this dosage will vary according to the tissue site of injection. The appropriate and effective dosage of nucleic acid sequence can readily be determined by those of ordinary skill in the art and may depend on the condition being treated and the route of administration. The preferred route of administration is by the parenteral route of injection into the interstitial space of tissues. However, other parenteral routes may also be used, such as, inhalation of an aerosol formulation particularly for delivery to lungs or bronchial tissues, throat or mucous membranes of the nose. In addition, naked polynucleotide constructs can be delivered to arteries during angioplasty by the catheter used in the procedure.

The dose response effects of injected polynucleotide in muscle *in vivo* is determined as follows. Suitable template DNA for production of mRNA coding for polypeptide of the present invention is prepared in accordance with a standard recombinant DNA methodology. The template DNA, which may be either circular or linear, is either used as naked DNA or complexed with liposomes. The quadriceps muscles of mice are then injected with various amounts of the template DNA.

Five to six week old female and male Balb/C mice are anesthetized by intraperitoneal injection with 0.3 ml of 2.5% Avertin. A 1.5 cm incision is made on the anterior thigh, and the quadriceps muscle is directly visualized. The template DNA is injected in 0.1 ml of carrier in a 1 cc syringe through a 27 gauge needle over one minute,

approximately 0.5 cm from the distal insertion site of the muscle into the knee and about 0.2 cm deep. A suture is placed over the injection site for future localization, and the skin is closed with stainless steel clips.

After an appropriate incubation time (e. g., 7 days) muscle extracts are prepared by excising the entire quadriceps. Every fifth 15 um cross-section of the individual quadriceps muscles is histochemically stained for protein expression. A time course for protein expression may be done in a similar fashion except that quadriceps from different mice are harvested at different times. Persistence of DNA in muscle following injection may be determined by Southern blot analysis after preparing total cellular DNA and HIRT supernatants from injected and control mice.

The results of the above experimentation in mice can be use to extrapolate proper dosages and other treatment parameters in humans and other animals using naked DNA.

Example 12: Transgenic Animals.

The polypeptides of the invention can also be expressed in transgenic animals. Animals of any species, including, but not limited to, mice, rats, rabbits, hamsters, guinea pigs, pigs, micro-pigs, goats, sheep, cows and non-human primates, e. g., baboons, monkeys, and chimpanzees may be used to generate transgenic animals. In a specific embodiment, techniques described herein or otherwise known in the art, are used to express polypeptides of the invention in humans, as part of a gene therapy protocol.

Any technique known in the art may be used to introduce the transgene (I. e., polynucleotides of the invention) into animals to produce the founder lines of transgenic animals. Such techniques include, but are not limited to, pronuclear microinjection (Paterson et al.,

Appl. Microbiol. Biotechnol. 40: 691-698 (1994); Carver et al., Biotechnology (NY) 11: 1263-1270 (1993); Wright et al., Biotechnology (NY) 9: 830-834 (1991); and Hoppe et al., U. S. Pat. No. 4,873,191 (1989)); retrovirus mediated gene transfer into germ lines (Van der Putten et al., Proc. Natl. Acad. Sci., USA 82: 6148-6152 (1985)), blastocysts or embryos; gene targeting in embryonic stem cells (Thompson et al., Cell 56: 313-321 (1989)); electroporation of cells or embryos (Lo, 1983, Mol Cell. Biol. 3: 1803-1814 (1983)); introduction of the polynucleotides of the invention using a gene gun (see, e. g., Ulmer et al., Science 259: 1745 (1993); introducing nucleic acid constructs into embryonic pleuripotent stem cells and transferring the stem cells back into the blastocyst; and sperm mediated gene transfer (Lavitrano et al., Cell 57: 717-723 (1989); etc. For a review of such techniques, see Gordon,"Transgenic Animals," Intl. Rev. Cytol. 115: 171-229 (1989), which is incorporated by reference herein in its entirety.

Any technique known in the art may be used to produce transgenic clones containing polynucleotides of the invention, for example, nuclear transfer into enucleated oocytes of nuclei from cultured embryonic, fetal, or adult cells induced to quiescence (Campell et al., Nature 380: 64-66 (1996); Wilmut et al., Nature 385: 810813 (1997)).

The present invention provides for transgenic animals that carry the transgene in all their cells, as well as animals which carry the transgene in some, but not all their cells, I. e., mosaic animals or chimeric. The transgene may be integrated as a single transgene or as multiple copies such as in concatamers, e. g., head-to-head tandems or head-to-tail tandems. The transgene may also be selectively introduced into and activated in a particular cell type by following, for example, the teaching of Lasko et al. (Lasko et al., Proc. Natl. Acad. Sci. USA 89: 6232-

6236 (1992)). The regulatory sequences required for such a cell-type specific activation will depend upon the particular cell type of interest, and will be apparent to those of skill in the art. When it is desired that the polynucleotide transgene be integrated into the chromosomal site of the endogenous gene, gene targeting is preferred. Briefly, when such a technique is to be utilized, vectors containing some nucleotide sequences homologous to the endogenous gene are designed for the purpose of integrating, via homologous recombination with chromosomal sequences, into and disrupting the function of the nucleotide sequence of the endogenous gene. The transgene may also be selectively introduced into a particular cell type, thus inactivating the endogenous gene in only that cell type, by following, for example, the teaching of Gu et al. (Gu et al., Science 265: 103-106 (1994)). The regulatory sequences required for such a cell-type specific inactivation will depend upon the particular cell type of interest, and will be apparent to those of skill in the art.

Once transgenic animals have been generated, the expression of the recombinant gene may be assayed utilizing standard techniques. Initial screening may be accomplished by Southern blot analysis or PCR techniques to analyze animal tissues to verify that integration of the transgene has taken place. The level of mRNA expression of the transgene in the tissues of the transgenic animals may also be assessed using techniques which include, but are not limited to, Northern blot analysis of tissue samples obtained from the animal, in situ hybridization analysis, and reverse transcriptase-PCR (rt-PCR). Samples of transgenic gene-expressing tissue may also be evaluated immunocytochemically or immunohistochemically using antibodies specific for the transgene product.

Once the founder animals are produced, they may be

bred, inbred, outbred, or crossbred to produce colonies of the particular animal. Examples of such breeding strategies include, but are not limited to: outbreeding of founder animals with more than one integration site in order to establish separate lines; inbreeding of separate lines in order to produce compound transgenics that express the transgene at higher levels because of the effects of additive expression of each transgene; crossing of heterozygous transgenic animals to produce animals homozygous for a given integration site in order to both augment expression and eliminate the need for screening of animals by DNA analysis; crossing of separate homozygous lines to produce compound heterozygous or homozygous lines; and breeding to place the transgene on a distinct background that is appropriate for an experimental model of interest.

Transgenic animals of the invention have uses which include, but are not limited to, animal model systems useful in elaborating the biological function of polypeptides of the present invention, studying conditions and/or disorders associated with aberrant expression, and in screening for compounds effective in ameliorating such conditions and/or disorders.

Example 13: Knock-Out Animals

Endogenous gene expression can also be reduced by inactivating or"knocking out"the gene and/or its promoter using targeted homologous recombination. (E. g., see Smithies et al., Nature 317: 230-234 (1985); Thomas & Capecchi, Cell 51: 503512 (1987); Thompson et al., Cell 5: 313-321 (1989); each of which is incorporated by reference herein in its entirety). For example, a mutant, non-functional polynucleotide of the invention (or a completely unrelated DNA sequence) flanked by DNA homologous to the endogenous polynucleotide sequence (either the coding regions or regulatory regions of the gene) can be used,

with or without a selectable marker and/or a negative selectable marker, to transfect cells that express polypeptides of the invention *in vivo*. In another embodiment, techniques known in the art are used to generate knockouts in cells that contain, but do not express the gene of interest. Insertion of the DNA construct, via targeted homologous recombination, results in inactivation of the targeted gene. Such approaches are particularly suited in research and agricultural fields where modifications to embryonic stem cells can be used to generate animal offspring with an inactive targeted gene (e. g., see Thomas & Capecchi 1987 and Thompson 1989, supra). However this approach can be routinely adapted for use in humans provided the recombinant DNA constructs are directly administered or targeted to the required site *in vivo* using appropriate viral vectors that will be apparent to those of skill in the art.

In further embodiments of the invention, cells that are genetically engineered to express the polypeptides of the invention, or alternatively, that are genetically engineered not to express the polypeptides of the invention (e. g., knockouts) are administered to a patient *in vivo*. Such cells may be obtained from the patient (I. e., animal, including human) or an MHC compatible donor and can include, but are not limited to fibroblasts, bone marrow cells, blood cells (e. g., lymphocytes), adipocytes, muscle cells, endothelial cells etc. The cells are genetically engineered *in vitro* using recombinant DNA techniques to introduce the coding sequence of polypeptides of the invention into the cells, or alternatively, to disrupt the coding sequence and/or endogenous regulatory sequence associated with the polypeptides of the invention, e. g., by transduction (using viral vectors, and preferably vectors that integrate the transgene into the cell genome) or transfection procedures, including, but not limited to,

the use of plasmids, cosmids, YACs, naked DNA, electroporation, liposomes, etc.

The coding sequence of the polypeptides of the invention can be placed under the control of a strong constitutive or inducible promoter or promoter/enhancer to achieve expression, and preferably secretion, of the polypeptides of the invention. The engineered cells which express and preferably secrete the polypeptides of the invention can be introduced into the patient systemically, e. g., in the circulation, or intraperitoneally.

Alternatively, the cells can be incorporated into a matrix and implanted in the body, e. g., genetically engineered fibroblasts can be implanted as part of a skin graft; genetically engineered endothelial cells can be implanted as part of a lymphatic or vascular graft. (See, for example, Anderson et al. U. S. Patent No. 5,399,349; and Mulligan & Wilson, U. S. Patent No. 5,460,959 each of which is incorporated by reference herein in its entirety).

When the cells to be administered are non-autologous or non-MHC compatible cells, they can be administered using well known techniques which prevent the development of a host immune response against the introduced cells. For example, the cells may be introduced in an encapsulated form which, while allowing for an exchange of components with the immediate extracellular environment, does not allow the introduced cells to be recognized by the host immune system.

Transgenic and"knock-out"animals of the invention have uses which include, but are not limited to, animal model systems useful in elaborating the biological function of polypeptides of the present invention, studying conditions and/or disorders associated with aberrant expression, and in screening for compounds effective in ameliorating such conditions and/or disorders.

It will be clear that the invention may be practiced

otherwise than as particularly described in the foregoing description and examples. Numerous modifications and variations of the present invention are possible in light of the above teachings and, therefore, are within the scope of the appended claims.

The entire disclosure of each document cited (including patents, patent applications, journal articles, abstracts, laboratory manuals, books, or other disclosures) in the Background of the Invention, Detailed Description, and Examples is hereby incorporated herein by reference. Further, the hard copy of the sequence listing submitted herewith and the corresponding computer readable form are both incorporated herein by reference in their entireties.

What is claimed is:

1. An LSG comprising:

(a) a polynucleotide of SEQ ID NO:1, 2, 5, or 6 or a variant thereof;

(b) a protein expressed by a polynucleotide of SEQ ID NO:1, 2, 5 or 6, or a variant thereof; or

(c) a polynucleotide which is capable of hybridizing under stringent conditions to the antisense sequence of SEQ ID NO: 1, 2, 5 or 6.

2. The LSG of claim 1 comprising a protein of SEQ ID NO: 3 or 4.

3. A method for diagnosing the presence of lung cancer in a patient comprising:

(a) determining levels of an LSG of claim 1 in cells, tissues or bodily fluids in a patient; and

(b) comparing the determined levels of LSG with levels of LSG in cells, tissues or bodily fluids from a normal human control, wherein a change in determined levels of LSG in said patient versus normal human control is associated with the presence of lung cancer.

4. A method of diagnosing metastases of lung cancer in a patient comprising:

(a) identifying a patient having lung cancer that is not known to have metastasized;

(b) determining levels of an LSG of claim 1 in a sample of cells, tissues, or bodily fluid from said patient; and

(c) comparing the determined LSG levels with levels of LSG in cells, tissue, or bodily fluid of a normal human control, wherein an increase in determined LSG levels in the patient versus the normal human control is associated with a cancer which has metastasized.

5. A method of staging lung cancer in a patient having lung cancer comprising:

(a) identifying a patient having lung cancer;

(b) determining levels of an LSG of claim 1 in a sample of cells, tissue, or bodily fluid from said patient; and

(c) comparing determined LSG levels with levels of LSG in cells, tissues, or bodily fluid of a normal human control, wherein an increase in determined LSG levels in said patient versus the normal human control is associated with a cancer which is progressing and a decrease in the determined LSG levels is associated with a cancer which is regressing or in remission.

6. A method of monitoring lung cancer in a patient for the onset of metastasis comprising:

(a) identifying a patient having lung cancer that is not known to have metastasized;

(b) periodically determining levels of an LSG of claim 1 in samples of cells, tissues, or bodily fluid from said patient; and

(c) comparing the periodically determined LSG levels with levels of LSG in cells, tissues, or bodily fluid of a normal human control, wherein an increase in any one of the periodically determined LSG levels in the patient versus the normal human control is associated with a cancer which has metastasized.

7. A method of monitoring a change in stage of lung cancer in a patient comprising:

(a) identifying a patient having lung cancer;

(b) periodically determining levels of an LSG of claim 1 in cells, tissues, or bodily fluid from said patient; and

(c) comparing the periodically determined LSG levels

with levels of LSG in cells, tissues, or bodily fluid of a normal human control, wherein an increase in any one of the periodically determined LSG levels in the patient versus the normal human control is associated with a cancer which is progressing in stage and a decrease is associated with a cancer which is regressing in stage or in remission.

8. A method of identifying potential therapeutic agents for use in imaging and treating lung cancer comprising screening molecules for an ability to bind to or decrease expression of an LSG of claim 1 relative to the LSG in the absence of the agent wherein the ability of a molecule to bind to the LSG or decrease expression of the LSG is indicative of the molecule being useful in imaging and treating lung cancer.

9. An antibody which specifically binds a polypeptide encoded by an LSG of claim 1.

10. A method of imaging lung cancer in a patient comprising administering to the patient an antibody of claim 9.

11. The method of claim 10 wherein said antibody is labeled with paramagnetic ions or a radioisotope.

12. A method of treating lung cancer in a patient comprising administering to the patient a molecule which downregulates expression or activity of an LSG of claim 1.

13. A method of inducing an immune response against a target cell expressing an LSG of claim 1 comprising delivering to a human patient an immunogenically stimulatory amount of an LSG protein so that an immune response is mounted against the target cell.

14. The method of claim 13 wherein the LSG protein comprises SEQ ID NO:3 or 4.

15. A vaccine for treating lung cancer comprising an LSG of claim 1.

SEQUENCE LISTING

<110> diaDexus, Inc.
Chen, Sei Yu
Sun, Yongming
Macina, Roberto

<120> Method of Diagnosing, Monitoring, Staging, Imaging and Treating Lung Cancer

<130> DEX-0139

<140>
<141>

<150> 60/183,188
<151> 2000-02-17

<160> 11

<170> PatentIn Ver. 2.1

<210> 1
<211> 1680
<212> DNA
<213> Homo sapiens

<400> 1
ggtgtgcagg atataaggtt ggacttccag acccactgcc cgggagagga sgggagcggg 60
ccgaggactc cagcgtgccc aggtctggca tcctgcactt gctgccctct gacacctggg 120
aagatggccg gcccgtggac cttcaccctt ctctgtggtt tgctggcagc caccttgatc 180
caagccaccc tcagtcccac tgcagttctc atcctcggcc caaaagtcat caaagaaaag 240
ctgacacagg agctgaagga ccacaacgcc accagcatcc tgcagcagct gccgctgctc 300
agtgccatgc gggaaaagcc agccggagga tccctgtgct gggcagcctg gtgaacaccg 360
tcctgaagca catcatctgg ctgaaggtca tcacagctaa catcctccag ctgcaggtga 420
agccctcggc caatgaccag gagctgctag tcaagatccc cctggacatg gtggctggat 480
tcaacacgcc cctggtcaag accatcgtgg agttccacat gacgactgag gcccaagcca 540
ccatccgcat ggacaccagt gcaagtggcc ccacccgcct ggtcctcagt gactgtgcca 600
ccagccatgg gagcctgcgc atccaactgc tgcataagct ctccttcctg gtgaacgcct 660
tagctaagca ggtcatgaac ctcctagtgc catccctgcc caatctagtg aaaaaccagc 720
tgtgtcccgt gatcgaggct tccttcaatg gcatgtatgc agacctcctg cagctggtga 780
aggtgcccat ttccctcagc attgaccgtc tggagtttga ccttctgtat cctgccatca 840
agggtgacac cattcagctc tacctggggg ccaagttgtt ggactcacag ggaaaggtga 900
ccaagtggtt caataactct gcagcttccc tgacaatgcc caccctggac aacatcccgt 960
tcagcctcat cgtgagtcag gacgtggtga aagctgcagt ggctgctgtg ctctctccag 1020
aagaattcat ggtcctgttg gactctgtgc ttcctgagag tgcccatcgg ctgaagtcaa 1080
gcatcgggct gatcaatgaa aaggctgcag ataagctggg atctacccag atcgtgaaga 1140
tcctaactca ggacactccc gagtttttta tagaccaagg ccatgccaag gtggcccaac 1200

tgatcgtgct ggaagtgttt ccctccagtg aagccctccg ccctttgttc accctgggca 1260
tcgaagccag ctcggaagct cagttttaca ccaaaggtga ccaacttata ctcaacttga 1320
ataacatcag ctctgatcgg atccagctga tgaactctgg gattggctgg ttccaacctg 1380
atgttctgaa aaacatcatc actgagatca tccactccat cctgctgccg aaccagaatg 1440
gcaaattaag atctggggtc ccagtgtcat tggtgaaggc cttgggattc gaggcagctg 1500
agtcctcact gaccaaggat gcccttgtgc ttactccagc ctccttgtgg aaacccagct 1560
ctcctgtctc ccagtgaaga cttggatggc agccatcagg gaaggctggg tcccagctgg 1620
gagtatgggt gtgagctcta tagaccatcc ctctctgcaa tcaataaaca cttgcctgtg 1680

<210> 2
<211> 895
<212> DNA
<213> Homo sapiens

<400> 2
ctaatctgtt acgtaacagc aagacagcgt cacctcacct gttctcgccc tcaaatggga 60
acgctggcct gggactaaag catagaccac caggctgagt atcctgacct gagtcatccc 120
cagggatcag gagcctccag cagggaacct tccattatat tcttcaagca acttacagct 180
gcaccgacag ttgcgatgaa agttctaatc tcttccctcc tcctgttgct gccactaatg 240
ctgatgtcca tggtctctag cagcctgaat ccaggggtcg ccagaggcca cagggaccga 300
ggccaggctt ctaggagatg gctccaggaa ggcggccaag aatgtgagtg caaagattgg 360
ttcctgagag ccccgagaag aaaattcatg acagtgtctg ggctgccaaa gaagcagtgc 420
ccctgtgatc atttcaaggg caatgtgaag aaaacaagac accaaaggca ccacagaaag 480
ccaaacaagc attccagagc ctgccagcaa tttctcaaac aatgtcagct aagaagcttt 540
gctctgcctt tgtaggagct ctgagcgccc actcttccaa ttaaacattc tcagccaaga 600
agacagtgag cacacctacc agacactctt cttctcccac ctcactctcc cactgtaccc 660
acccctaaat cattccagtg ctctcaaaaa gcatgttttt caagatcatt ttgtttgttg 720
ctctctctag tgtcttcttc tctcgtcagt cttagcctgt gccctcccct tacccaggct 780
taggcttaat tacctgaaag attccaggaa actgtagctt cctagctagt gtcatttaac 840
cttaaatgca atcaggaaag tagcaaacag aagtcaataa atatttttaa atgtc 895

<210> 3
<211> 484
<212> PRT
<213> Homo sapiens

<400> 3
Met Ala Gly Pro Trp Thr Phe Thr Leu Leu Cys Gly Leu Leu Ala Ala
1 5 10 15

Thr Leu Ile Gln Ala Thr Leu Ser Pro Thr Ala Val Leu Ile Leu Gly
20 25 30

Pro Lys Val Ile Lys Glu Lys Leu Thr Gln Glu Leu Lys Asp His Asn
35 40 45

```
Ala Thr Ser Ile Leu Gln Gln Leu Pro Leu Leu Ser Ala Met Arg Glu
     50                  55                  60

Lys Pro Ala Gly Gly Ile Pro Val Leu Gly Ser Leu Val Asn Thr Val
 65                  70                  75                  80

Leu Lys His Val Ile Trp Leu Lys Val Ile Thr Ala Asn Ile Leu Gln
                 85                  90                  95

Leu Gln Val Lys Pro Ser Ala Asn Asp Gln Glu Leu Leu Val Lys Ile
            100                 105                 110

Pro Leu Asp Met Val Ala Gly Phe Asn Thr Pro Leu Val Lys Thr Ile
        115                 120                 125

Val Glu Phe His Met Thr Thr Glu Ala Gln Ala Thr Ile Arg Met Asp
    130                 135                 140

Thr Ser Ala Ser Gly Pro Thr Arg Leu Val Leu Ser Asp Cys Ala Thr
145                 150                 155                 160

Ser His Gly Ser Leu Arg Ile Gln Leu Leu His Lys Leu Ser Phe Leu
                165                 170                 175

Val Asn Ala Leu Ala Lys Gln Val Met Asn Leu Leu Val Pro Ser Leu
            180                 185                 190

Pro Asn Leu Val Lys Asn Gln Leu Cys Pro Val Ile Glu Ala Ser Phe
        195                 200                 205

Asn Gly Met Tyr Ala Asp Leu Leu Gln Leu Val Lys Val Pro Ile Ser
    210                 215                 220

Leu Ser Ile Asp Arg Leu Glu Phe Asp Leu Leu Tyr Pro Ala Ile Lys
225                 230                 235                 240

Gly Asp Thr Ile Gln Leu Tyr Leu Gly Ala Lys Leu Leu Asp Ser Gln
                245                 250                 255

Gly Lys Val Thr Lys Trp Phe Asn Asn Ser Ala Ala Ser Leu Thr Met
            260                 265                 270

Pro Thr Leu Asp Asn Ile Pro Phe Ser Leu Ile Val Ser Gln Asp Val
        275                 280                 285

Val Lys Ala Ala Val Ala Ala Val Leu Ser Pro Glu Glu Phe Met Val
    290                 295                 300
```

```
Leu Leu Asp Ser Val Leu Pro Glu Ser Ala His Arg Leu Lys Ser Ser
305                 310                 315                 320

Ile Gly Leu Ile Asn Glu Lys Ala Ala Asp Lys Leu Gly Ser Thr Gln
                325                 330                 335

Ile Val Lys Ile Leu Thr Gln Asp Thr Pro Glu Phe Phe Ile Asp Gln
            340                 345                 350

Gly His Ala Lys Val Ala Gln Leu Ile Val Leu Glu Val Phe Pro Ser
        355                 360                 365

Ser Glu Ala Leu Arg Pro Leu Phe Thr Leu Gly Ile Glu Ala Ser Ser
    370                 375                 380

Glu Ala Gln Phe Tyr Thr Lys Gly Asp Gln Leu Ile Leu Asn Leu Asn
385                 390                 395                 400

Asn Ile Ser Ser Asp Arg Ile Gln Leu Met Asn Ser Gly Ile Gly Trp
                405                 410                 415

Phe Gln Pro Asp Val Leu Lys Asn Ile Ile Thr Glu Ile Ile His Ser
            420                 425                 430

Ile Leu Leu Pro Asn Gln Asn Gly Lys Leu Arg Ser Gly Val Pro Val
        435                 440                 445

Ser Leu Val Lys Ala Leu Gly Phe Glu Ala Ala Glu Ser Ser Leu Thr
    450                 455                 460

Lys Asp Ala Leu Val Leu Thr Pro Ala Ser Leu Trp Lys Pro Ser Ser
465                 470                 475                 480

Pro Val Ser Gln
```

<210> 4
<211> 119
<212> PRT
<213> Homo sapiens

<400> 4

```
Met Lys Val Leu Ile Ser Ser Leu Leu Leu Leu Leu Pro Leu Met Leu
  1               5                  10                  15

Met Ser Met Val Ser Ser Ser Leu Asn Pro Gly Val Ala Arg Gly His
             20                  25                  30
```

```
Arg Asp Arg Gly Gln Ala Ser Arg Arg Trp Leu Gln Glu Gly Gly Gln
         35                  40                  45

Glu Cys Glu Cys Lys Asp Trp Phe Leu Arg Ala Pro Arg Arg Lys Phe
     50                  55                  60

Met Thr Val Ser Gly Leu Pro Lys Lys Gln Cys Pro Cys Asp His Phe
 65                  70                  75                  80

Lys Gly Asn Val Lys Lys Thr Arg His Gln Arg His His Arg Lys Pro
                 85                  90                  95

Asn Lys His Ser Arg Ala Cys Gln Gln Phe Leu Lys Gln Cys Gln Leu
            100                 105                 110

Arg Ser Phe Ala Leu Pro Leu
        115
```

<210> 5
<211> 2036
<212> DNA
<213> Homo sapiens

<400> 5
```
gggggaggcc tctgcctaaa gacaaagcct gtgctggggt gtgcaggata taaggttgga 60
cttccagacc cactgcccgg gagaggagag aggagcgggc cgaggactcc agcgtgccca 120
ggtctggcat cctgcacttg ctgccctctg acacctggga agatggccgg cccgtggacc 180
ttcacccttc tctgtggttt gctggcagcc accttgatcc aagccaccct cagtcccact 240
gcagttctca tcctcggccc aaaagtcatc aaagaaaagc tgacacagga gctgaaggac 300
cacaacgcca ccagcatcct gcagcagctg ccgctgctca gtgccatgcg ggaaaagcca 360
gccggaggca tccctgtgct gggcagcctg gtgaacaccg tcctgaagca crtcatctgg 420
ctgaaggtca tcacagctaa catcctccag ctgcaggtga agccctcggc caatgaccag 480
gagctgctag tcaagatccc cctggacatg gtggctggat tcaacacgcc cctggtcaag 540
accatcgtgg agttccacat gacgactgag gcccaagcca ccatccgcat ggacaccagt 600
gcaagtggcc ccacccgcct ggtcctcagt gactgtgcca ccagccatgg gagcctgcgc 660
atccaactgc tgcataagct ctccttcctg gtgaacgcct tagctaagca ggtcatgaac 720
ctcctagtgc catccctgcc caatctagtg aaaaaccagc tgtgtcccgt gatcgaggct 780
tccttcaatg gcatgtatgc agacctcctg cagctggtga aggtgcccat ttccctcagc 840
attgaccgtc tggagtttga ccttctgtat cctgccatca agggtgacac cattcagctc 900
tacctggggg ccaagttgtt ggactcacag ggaaaggtga ccaagtggtt caataactct 960
gcagcttccc tgacaatgcc caccctggac aacatcccgt tcagcctcat cgtgagtcag 1020
gacgtggtga aagctgcagt ggctgctgtg ctctctccag aagaattcat ggtcctgttg 1080
gactctgtgc ttcctgagag tgcccatcgg ctgaagtcaa gcatcgggct gatcaatgaa 1140
aaggctgcag ataagctggg atctacccag atcgtgaaga tcctaactca ggacactccc 1200
gagtttttta tagaccaagg ccatgccaag gtggcccaac tgatcgtgct ggaagtgttt 1260
ccctccagtg aagccctccg ccctttgttc accctgggca tcgaagccag ctcggaagct 1320
```

cagttttaca ccaaaggtga ccaacttata ctcaacttga ataacatcag ctctgatcgg 1380
atccagctga tgaactctgg gattggctgg ttccaacctg atgttctgaa aaacatcatc 1440
actgagatca tccactccat cctgctgccg aaccagaatg gcaaattaag atctggggtc 1500
ccagtgtcat tggtgaaggc cttgggattc gaggcagctg agtcctcact gaccaaggat 1560
gcccttgtgc ttactccagc ctccttgtgg aaacccagct ctcctgtctc ccagtgaaga 1620
cttggatggc agccatcagg gaaggctggg tcccagctgg gagtatgggt gtgagctcta 1680
tagaccatcc ctctctgcaa tcaataaaca cttgcctgtg aaaaaaaaaa aaaaaaaaaa 1740
aaaaaaaaaa ctcgactcta gattgcggcc gcggtcatag ctgtttcctg aacagatccc 1800
gggtggcatc cctgtgaccc ctccccagtg cctctcctgg ccctggaagt tgccactcca 1860
gtgcccacca gccttgtcct aataaaatta agttgcatca ttttgtctga ctaggtgtcc 1920
ttctataata ttatggggtg gaggggggtg gtatggagca aggggcaagt tgggaagaca 1980
acctgtaggg cctgcggggt ctattgggaa ccaagctgga gtgcagtggc acaatc 2036

<210> 6
<211> 1171
<212> DNA
<213> Homo sapiens

<400> 6
gattcccata aagcacatgg tctaatctgt tacgtaacag caagacagcg tcacctcacc 60
tgttctcgcc ctcaaatggg aacgctggcc tgggactaaa gcatagacca ccaggctgag 120
tatcctgacc tgagtcatcc ccagggatca ggagcctcca gcagggaacc ttccattata 180
ttcttcaagc aacttacagc tgcaccgaca gttgcgatga aagttctaat ctcttccctc 240
ctcctgttgc tgccactaat gctgatgtcc atggtctcta gcagcctgaa tccaggggtc 300
gccagaggcc acagggaccg aggccaggct tctaggagat ggctccagga aggcggccaa 360
gaatgtgagt gcaaagattg gttcctgaga gccccgagaa gaaaattcat gacagtgtct 420
gggctgccaa agaagcagtg cccctgtgat catttcaagg gcaatgtgaa gaaaacaaga 480
caccaaaggc accacagaaa gccaaacaag cattccagag cctgccagca atttctcaaa 540
caatgtcagc taagaagctt tgctctgcct ttgtaggagc tctgagcgcc cactcttcca 600
attaaacatt ctcagccaag aagacagtga gcacacctac cagacactct tcttctccca 660
cctcactctc ccactgtacc cacccctaaa tcattccagt gctctcaaaa agcatgtttt 720
tcaagatcat tttgtttgtt gctctctcta gtgtcttctt ctctcgtcag tcttagcctg 780
tgccctcccc ttacccaggc ttaggcttaa ttacctgaaa gattccagga aactgtagct 840
tcctagctag tgtcatttaa ccttaaatgc aatcaggaaa gtagcaaaca gaagtcaata 900
aatattttta aatgtcacag atcaaaattg tttccttcaa atggggtctg ccaattcaca 960
accagatgac ccattttacc ctattcactg cagactgaat ccagattcta cacatactta 1020
tccccaccaa gaccctcact ctgtctccat tggcctactt gttcatcttt cactcattcg 1080
acaaatcttt ttgaggtaag agcgaggtgg gacaaaaaaa aaaagcatac caatgaacca 1140
gacccggtct tattaaagat aatataggtt t 1171

<210> 7
<211> 18
<212> DNA
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

<400> 7
ggcgttgtgg tccttcag 18

<210> 8
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 8
acttgctgcc ctctgacac 19

<210> 9
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 9
atcacagggg cactgcttct 20

<210> 10
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

<400> 10
gaccgaggcc aggcttcta 19

<210> 11
<211> 733
<212> DNA
<213> Homo sapiens

<400> 11

```
gggatccgga gcccaaatct tctgacaaaa ctcacacatg cccaccgtgc ccagcacctg 60
aattcgaggg tgcaccgtca gtcttcctct tccccccaaa acccaaggac accctcatga 120
tctcccggac tcctgaggtc acatgcgtgg tggtggacgt aagccacgaa gaccctgagg 180
tcaagttcaa ctggtacgtg gacggcgtgg aggtgcataa tgccaagaca aagccgcggg 240
aggagcagta caacagcacg taccgtgtgg tcagcgtcct caccgtcctg caccaggact 300
ggctgaatgg caaggagtac aagtgcaagg tctccaacaa agccctccca acccccatcg 360
agaaaaccat ctccaaagcc aaagggcagc cccgagaacc acaggtgtac accctgcccc 420
catcccggga tgagctgacc aagaaccagg tcagcctgac ctgcctggtc aaaggcttct 480
atccaagcga catcgccgtg gagtgggaga gcaatgggca gccggagaac aactacaaga 540
ccacgcctcc cgtgctggac tccgacggct ccttcttcct ctacagcaag ctcaccgtgg 600
acaagagcag gtggcagcag gggaacgtct tctcatgctc cgtgatgcat gaggctctgc 660
acaaccacta cacgcagaag agcctctccc tgtctccggg taaatgagtg cgacggccgc 720
gactctagag gat                                                   733
```

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
23 August 2001 (23.08.2001)

PCT

(10) International Publication Number
WO 01/061055 A3

(51) International Patent Classification[7]: C12Q 1/68

(21) International Application Number: PCT/US01/05674

(22) International Filing Date: 20 February 2001 (20.02.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/183,188 17 February 2000 (17.02.2000) US

(71) Applicant: DIADEXUS, INC. [US/US]; 343 Oyster Point Boulevard, San Francisco, CA 94080 (US).

(72) Inventors: CHEN, Sei-Yu; 160 Mira Street, Foster City, CA 94404 (US). SUN, Yongming; Apartment 260, 869 S. Winchester Boulevard, Apartment 260, San Jose, CA 95128 (US). MACINA, Roberto, A.; 4118 Crescendo Avenue, San Jose, CA 95136 (US).

(74) Agents: LICATA, Jane, Massey et al.; Licata & Tyrrell P.C., 66 E. Main Street, Marlton, NJ 08053 (US).

(81) Designated States *(national)*: AU, CA, JP.

(84) Designated States *(regional)*: European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR).

Published:
— *with international search report*

(88) Date of publication of the international search report:
29 August 2002

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 01/061055 A3

(54) Title: METHODS FOR DIAGNOSING, MONITORING, STAGING, IMAGING AND TREATING LUNG CANCER VIA LUNG CANCER SPECIFIC GENES

(57) Abstract: The invention relates to LSG polypeptides, polynucleotides encoding the polypeptides, methods for producing the polypeptides, in particular by expressing the polynucleotides, and agonists and antagonists of the polypeptides. The invention further relates to methods for utilizing such polynucleotides, polypeptides, agonists and antagonists for applications, which relate, in part, to research, diagnostic and clinical arts.

INTERNATIONAL SEARCH REPORT

	ıational Application No PCT/US 01/05674

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12Q1/68

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
EPO-Internal, WPI Data, PAJ, SEQUENCE SEARCH, MEDLINE, BIOSIS

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 99 60160 A (DIADEXUS LLC ;SUN YONGMING (US); YANG FEI (US); MACINA ROBERTO A () 25 November 1999 (1999-11-25) the whole document	1-15
X	WO 00 01821 A (INCYTE PHARMA INC ;PATTERSON CHANDRA (US); CORLEY NEIL C (US); YUE) 13 January 2000 (2000-01-13) page 67; claim 7; table 1	1,2,9,15
A	WO 99 34017 A (ABBOTT LAB) 8 July 1999 (1999-07-08) the whole document	1-15
A	US 5 939 265 A (ROBERTS-RAPP LISA ET AL) 17 August 1999 (1999-08-17) the whole document	1-15

☐ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
23 May 2002	29/05/2002

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Reuter, U

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.1

Although claims 10 and 11 are directed to a diagnostic method practised on the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.
Although claims 12-14 are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.

INTERNATIONAL SEARCH REPORT

ional Application No
CT/US 01/05674

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9960160 A	25-11-1999	CA 2328138 A1	25-11-1999
		EP 1082459 A1	14-03-2001
		WO 9960160 A1	25-11-1999
WO 0001821 A	13-01-2000	AU 4857999 A	24-01-2000
		EP 1092026 A2	18-04-2001
		WO 0001821 A2	13-01-2000
WO 9934017 A	08-07-1999	CA 2316017 A1	08-07-1999
		EP 1042506 A1	11-10-2000
		WO 9934017 A1	08-07-1999
		US 2002045165 A1	18-04-2002
US 5939265 A	17-08-1999	WO 9820143 A1	14-05-1998